

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number
WO 02/086443 A2

- (51) International Patent Classification⁷: **G01N**
- (21) International Application Number: **PCT/US02/12476**
- (22) International Filing Date: **18 April 2002 (18.04.2002)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:
- | | | |
|------------|-------------------------------|----|
| 60/284,770 | 18 April 2001 (18.04.2001) | US |
| 60/290,492 | 10 May 2001 (10.05.2001) | US |
| 60/339,245 | 9 November 2001 (09.11.2001) | US |
| 60/350,666 | 13 November 2001 (13.11.2001) | US |
| 60/334,370 | 29 November 2001 (29.11.2001) | US |
| 60/372,246 | 12 April 2002 (12.04.2002) | US |
- (71) Applicant (*for all designated States except US*): **EOS BIOTECHNOLOGY, INC.** [US/US]; 225A Gateway Boulevard, South San Francisco, CA 94080 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **AZIZ, Natasha** [US/US]; 411 California Avenue, Palo Alto, CA 94306 (US). **MURRAY, Richard** [US/US]; 22643 Woodbridge Court, Cupertino, CA 95014 (US).
- (74) Agents: **BASTIAN, Kevin, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111-3834 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Declaration under Rule 4.17:**
— *of inventorship (Rule 4.17(iv)) for US only*
- Published:**
— *without international search report and to be republished upon receipt of that report*
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF LUNG CANCER

(57) Abstract: Described herein are methods and compositions that can be used for diagnosis and treatment of lung cancer and similar pathologies. Also described herein are methods that can be used to identify modulators of lung cancer and similar pathologies.

WO 02/086443 A2

METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS
OF SCREENING FOR MODULATORS OF LUNG CANCER

5

CROSS-REFERENCES TO RELATED APPLICATIONS

This application is related to USSN 60/284,770, filed April 18, 2001; USSN 60/290,492, filed May 10, 2001; USSN 60/334,370, filed November 29, 2001; USSN 60/339,245, filed November 9, 2001; USSN 60/350,666, filed November 13, 2001; and
10 USSN 60/xxx,xxx, filed April 12, 2002 (Docket OMNI-002P); each of which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

The invention relates to the identification of nucleic acid and protein expression
15 profiles and nucleic acids, products, and antibodies thereto that are involved in lung cancer; and to the use of such expression profiles and compositions in diagnosis and therapy of lung cancer. The invention further relates to methods for identifying and using agents and/or targets that inhibit lung cancer or related conditions.

20

BACKGROUND OF THE INVENTION

Lung cancer is the second most commonly occurring cancer in the United States and is the leading cause of cancer-related death. It is estimated that there are over 160,000 new cases of lung cancer in the United States every year. Of those who are diagnosed with lung cancer, 86 percent will die within five years. Lung cancer is the most common visceral
25 cancer in men and accounts for nearly one third of all cancer deaths in both men and women. In fact, lung cancer accounts for 7% of all deaths, due to any cause, in both men and women.

Smoking is the primary cause of lung cancer, with more than 80% of lung cancers resulting from smoking. About 400 to 500 separate gaseous substances are present in the smoke of a non-filter cigarette. The most noteworthy substances include nitrogen oxides,
30 hydrogen cyanide, formaldehyde, benzene, and toluene. The particles present in cigarette smoke contain at least 3,500 individual compounds such as nicotine, tobacco alkaloids (nicotine, anatabine, anabasine), polycyclic aromatic hydrocarbons (e.g., benzo(a)pyrene, B(a)P), naphthalenes, aromatic amines, phenols, and tobacco-specific nitrosamines.

Tobacco-specific nitrosamines are formed during tobacco curing and processing, and are suspected of causing lung cancer in humans. In rodent studies, regardless of the where or how it is applied, the tobacco-specific nitrosamine known as NNK produces lung adenomas and lung adenocarcinomas. The tobacco-specific nitrosamine known as NNAL also produces lung adenocarcinomas in rodents.

Many of the chemicals found in cigarette smoke also affect the nonsmoker inhaling "secondhand" or sidestream smoke. Indeed, the smoke inhaled by non-smokers has a chemical composition similar to the smoke inhaled by smokers, but, importantly, the concentrations of the carcinogenic tobacco-specific nitrosamines are present in higher concentrations in second hand smoke. For this and other reasons, "passive smoking" is an important cause of lung cancer, causing as many as 3,000 lung cancer deaths in nonsmokers each year.

In addition to smoking, other factors thought to be causes of lung cancer include on-the-job exposure to carcinogens such as asbestos and uranium, exposure to chemical hazards such as radon, polycyclic aromatic hydrocarbons, chromium, nickel, and inorganic arsenic, genetic factors, and diet.

Histological classification of various lung cancers define the types of cancer that begin in the lung. See, e.g., Travis, et al. (1999) Histological Typing of Lung and Pleural Tumours (International Histological Classification of Tumours, No 1. Four major cell types make up more than 88% of all primary lung neoplasms. These are: squamous or epidermoid carcinoma, small cell (also called oat cell) carcinoma, adenocarcinoma, and large cell (also called large cell anaplastic) carcinoma. The remainder include undifferentiated carcinomas, carcinoids, bronchial gland tumors, and other rarer types. The various cell types have different natural histories and responses to therapy, and, thus, a correct histologic diagnosis is the first step of effective treatment.

Small cell lung cancer (SCLC) accounts for 18-25% of all lung cancers, and occurs less frequently than non-small cell lung cancers, and generally spread to distant organs more rapidly than non-small cell lung cancer. In general, at the time of presentation small cell lung cancers have already spread beyond the bounds where surgery and curative intent can be undertaken. However, if identified early enough, these cancers are often responsive to chemotherapy and thoracic radiation treatment.

Non-small cell lung cancers (NSCLC) are the more frequently occurring form of lung cancer. They comprise squamous cell carcinoma, adenocarcinoma, and large cell carcinoma

and account for more than 75% of all lung cancers. Non-small cell tumors that are localized at the time of presentation can sometimes be cured with surgery and/or radiotherapy, but usually are not identified until significant metastasis has occurred, which are typically not very responsive to surgical, chemotherapy, or radiation treatment..

5 The screening of asymptomatic persons at high risk for lung cancer has often proven ineffective. In general, only 5 to 15 percent of lung cancer patients have their disease detected while they are asymptomatic. Of course, early detection and treatment are critical factors in the fight against lung cancer. The average survival rate is 49% for those whose cancer is detected early, before the cancer has spread from the lung. Lung cancer often
10 spreads outside of the lung, and it may have spread to the bones or brain by the time it is diagnosed. While the prognosis may be better for lung cancers that are detected early, because of the lack ofv effective curative treatments, early detection does not necessarily alter the total death rate from lung cancer.

 Thus, methods for diagnosis and prognosis of lung cancer and effective treatment of
15 lung cancer would be desirable. Accordingly, provided herein are methods that can be used in diagnosis and prognosis of lung cancer. Further provided are methods that can be used to screen candidate therapeutic agents for the ability to modulate, e.g., treat, lung cancer. Additionally, provided herein are molecular targets and compositions for therapeutic intervention in lung disease and other metastatic cancers.

20

SUMMARY OF THE INVENTION

 The present invention provides nucleotide sequences of genes that are up- and down-regulated in lung cancer cells. Such genes are useful for diagnostic purposes, and also as targets for screening for therapeutic compounds that modulate lung cancer, such as
25 antibodies. The methods of detecting nucleic acids of the invention or their encoded proteins can be used for a number of purposes. Examples include early detection of lung cancers, monitoring and early detection of relapse following treatment of lung cancers, monitoring response to therapy of lung cancers, determining prognosis of lung cancers, directing therapy of lung cancers, selecting patients for postoperative chemotherapy or radiation therapy,
30 selecting therapy, determining tumor prognosis, treatment, or response to treatment, and early detection of precancerous lesions of the lung. Examples of benign or precancerous lesions include: atelectasis, emphysema, brochitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis (HP), interstitial pulmonary fibrosis (IPF), asthma, and

bronchiectasis. Other aspects of the invention will become apparent to the skilled artisan by the following description of the invention.

In one aspect, the present invention provides a method of detecting a lung cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16. Alternatively, the sample may be contacted with a specific binding reagent, e.g., antibody.

In one embodiment, the polynucleotide selectively hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-16. In another embodiment, the polynucleotide comprises a sequence as shown in Tables 1A-16.

In one embodiment, the biological sample is a tissue sample, or a body fluid. In another embodiment, the biological sample comprises isolated nucleic acids, e.g., mRNA.

In one embodiment, the polynucleotide is labeled, e.g., with a fluorescent label. In one embodiment, the polynucleotide is immobilized on a solid surface. In one embodiment, the patient is undergoing a therapeutic regimen to treat lung cancer. In another embodiment, the patient is suspected of having lung cancer. In one embodiment, the patient is a primate, e.g., a human.

In one embodiment, the method further comprises the step of amplifying nucleic acids before the step of contacting the biological sample with the polynucleotide.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated transcript in the biological sample by contacting the biological sample with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby monitoring the efficacy of the therapy. Or the sample may be evaluated for protein, e.g., contacting the sample with an antibody.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated transcript to a level of the lung cancer-associated transcript in a biological sample from the patient prior to, or earlier in, the therapeutic treatment. Or the sample may be evaluated for comparison of protein.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a

biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated antibody in the biological sample by contacting the biological sample with a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, wherein the polypeptide specifically binds to the lung cancer-associated antibody, thereby monitoring the efficacy of the therapy.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated antibody to a level of the lung cancer-associated antibody in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated polypeptide in the biological sample by contacting the biological sample with an antibody, wherein the antibody specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby monitoring the efficacy of the therapy.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated polypeptide to a level of the lung cancer-associated polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic treatment. In one aspect, the present invention provides an isolated nucleic acid molecule consisting of a polynucleotide sequence as shown in Tables 1A-16. In one embodiment, an expression vector or cell comprises the isolated nucleic acid. In one aspect, the present invention provides an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-16.

In another aspect, the present invention provides an antibody that specifically binds to an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-16. In one embodiment, the antibody is conjugated to an effector component, e.g., a fluorescent label, a radioisotope or a cytotoxic chemical. In one embodiment, the antibody is an antibody fragment. In another embodiment, the antibody is humanized.

In one aspect, the present invention provides a method of detecting lung cancer in a patient, the method comprising contacting a biological sample from the patient with an antibody or protein as described herein.

In another aspect, the present invention provides a method of detecting antibodies specific to a lung cancer gene in a patient, the method comprising contacting a biological sample from the patient with a polypeptide encoded by a nucleic acid comprises a sequence from Tables 1A-16.

5 In another aspect, the present invention provides a method for identifying a compound that modulates a lung cancer-associated polypeptide, the method comprising the steps of: (i) contacting the compound with a lung cancer-associated polypeptide, the polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16; and (ii) determining the functional effect of the
10 compound upon the polypeptide.

In one embodiment, the functional effect is a physical effect, an enzymatic effect, or a chemical effect. In one embodiment, the polypeptide is expressed in a eukaryotic host cell or cell membrane. In another embodiment, the polypeptide is recombinant. In one
15 embodiment, the functional effect is determined by measuring ligand binding to the polypeptide.

In another aspect, the present invention provides a method of inhibiting proliferation or another critical process of a lung cancer-associated cell to treat lung cancer in a patient, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified as described herein. In one embodiment, the compound is an
20 antibody.

In another aspect, the present invention provides a drug screening assay comprising the steps of: (i) administering a test compound to a mammal having lung cancer or a cell isolated therefrom; (ii) comparing the level of gene expression of a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables
25 1A-16 in a treated cell or mammal with the level of gene expression of the polynucleotide in a control cell or mammal, wherein a test compound that modulates the level of expression of the polynucleotide is a candidate for the treatment of lung cancer.

In one embodiment, the control is a mammal with lung cancer or a cell therefrom that has not been treated with the test compound. In another embodiment, the control is a normal
30 cell or mammal, or a non-malignant lung disease.

In another aspect, the present invention provides a method for treating a mammal having lung cancer comprising administering a compound identified by the assay described herein.

In another aspect, the present invention provides a pharmaceutical composition for treating a mammal having lung cancer, the composition comprising a compound identified by the assay described herein and a physiologically acceptable excipient.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the objects outlined above, the present invention provides novel methods for diagnosis and treatment of lung disease or cancer, as well as methods for screening for compositions which modulate lung cancer. "Treatment, monitoring, detection or modulation of lung disease or cancer" includes treatment, monitoring, detection, or modulation of lung disease in those patients who have lung disease (whether malignant or non-malignant, e.g., emphysema, bronchitis, or fibrosis) as well as patients with lung cancers in which gene expression from a gene in Tables 1A-16 is increased or decreased, indicating that the subject is more likely to have disease. In particular, while these targets are identified primarily from lung cancer samples, these same targets are likely to be similarly found in analyses of other medical conditions. These other conditions may result from similar pathological processes which affect similar tissues, e.g., lung cancer, small cell lung carcinoma (oat cell carcinoma), non-small cell carcinomas (e.g., squamous cell carcinoma, adenocarcinoma, large cell lung carcinoma, carcinoid, granulomatous), fibrosis (idiopathic pulmonary fibrosis (IPF), hypersensitivity pneumonitis (HP), interstitial pneumonitis, nonspecific idiopathic pneumonitis (NSIP)), chronic obstructive pulmonary disease (COPD, e.g., emphysema, chronic bronchitis), asthma, bronchiectasis, and esophageal cancer. See, e.g., the NCI webpage and USSN 60/347,349 and USSN 60/xxx,xxx (docket LFBR-001-1P, filed March 29, 2002), each of which is incorporated herein by reference. The treatment may be of lung cancer or related condition itself, or treatment of metastasis.

In particular, identification of markers selectively expressed on these cancers allows for use of that expression in diagnostic, prognostic, or therapeutic methods. As such, the invention defines various compositions, e.g., nucleic acids, polypeptides, antibodies, and small molecule agonists/antagonists, which will be useful to selectively identify those markers. For example, therapeutic methods may take the form of protein therapeutics which use the marker expression for selective localization or modulation of function (for those markers which have a causative disease effect), for vaccines, identification of binding partners, or antagonism, e.g., using antisense or RNAi. The markers may be useful for molecular characterization of subsets of lung diseases, which subsets may actually require

very different treatments. Moreover, the markers may also be important in related diseases to the specific cancers, e.g., which affect similar tissues in non-malignant diseases, or have similar mechanisms of induction/maintenance. Metastatic processes or characteristics may also be targeted. Diagnostic and prognostic uses are made available, e.g., to subset related but distinct diseases, or to determine treatment strategy. The detection methods may be based upon nucleic acid, e.g., PCR or hybridization techniques, or protein, e.g., ELISA, imaging, IHC, etc. The diagnosis may be qualitative or quantitative, and may detect increases or decreases in expression levels.

Tables 1A-16 provide unigene cluster identification numbers for the nucleotide sequence of genes that exhibit increased or decreased expression in lung cancer samples. The tables also provide an exemplar accession number that provides a nucleotide sequence that is part of the unigene cluster. In Table 1A, genes marked as "target 1" or "target 2" are particularly useful as therapeutic targets. Genes marked as "target 3" are particularly useful as diagnostic markers. Genes marked as "chron" are upregulated in chronically diseased lung (e.g., emphysema, bronchitis, fibrosis) relative to lung tumors and normal tissue. In certain analyses, the ratio for the "chron" category was determined using the 70th percentile of chronically diseased lung samples divided by the 90th percentile of normal lung samples. The ratio for the targets was determined using the 70th percentile of lung tumor samples divided by the 90th percentile of normal lung samples.

20

Definitions

The term "lung cancer protein" or "lung cancer polynucleotide" or "lung cancer-associated transcript" refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have a nucleotide sequence that has greater than about 60% nucleotide sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or greater nucleotide sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid sequence, or the complement thereof of Tables 1A-16 and conservatively modified variants thereof; or (4)

have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or greater amino sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acid, to an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or other mammal. A "lung cancer polypeptide" and a "lung cancer polynucleotide," include both naturally occurring or recombinant forms.

10 A "full length" lung cancer protein or nucleic acid refers to a lung cancer polypeptide or polynucleotide sequence, or a variant thereof, that contains the elements normally contained in one or more naturally occurring, wild type lung cancer polynucleotide or polypeptide sequences. The "full length" may be prior to, or after, various stages of post-translational processing or splicing, including alternative splicing.

15 "Biological sample" as used herein is a sample of biological tissue or fluid that contains nucleic acids or polypeptides, e.g., of a lung cancer protein, polynucleotide, or transcript. Such samples include, but are not limited to, tissue isolated from primates, e.g., humans, or rodents, e.g., mice, and rats. Biological samples may also include sections of tissues such as biopsy and autopsy samples, frozen sections taken for histologic purposes, archival materials, blood, plasma, serum, sputum, stool, tears, mucus, hair, skin, etc. Biological samples also include explants and primary and/or transformed cell cultures derived from patient tissues. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate, e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or other mammal; or a bird; reptile; 20 fish. Livestock and domestic animals are of interest.

"Providing a biological sample" means to obtain a biological sample for use in methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (e.g., isolated by another person, at another time, and/or for another purpose), or by performing the methods of the invention in vivo. Archival tissues or materials, having treatment or outcome history, will be particularly useful. 30

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the

same or have a specified percentage of amino acid residues or nucleotides that are the same (e.g., about 60% identity, preferably 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using, e.g., a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (see, e.g., NCBI web site <http://www.ncbi.nlm.nih.gov/BLAST/> or the like). Such sequences are then said to be "substantially identical." This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions and/or insertions, substitutions, and naturally occurring, e.g., polymorphic or allelic variants, and man-made variants. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of contiguous positions selected from the group consisting typically of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman (1988) Proc. Nat'l. Acad. Sci. USA 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer

Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel, et al. (eds. 1995 and supplements) Current Protocols in Molecular Biology.

Preferred examples of algorithms that are suitable for determining percent sequence identity and sequence similarity include the BLAST and BLAST 2.0 algorithms, which are described in Altschul, et al. (1977) Nuc. Acids Res. 25:3389-3402 and Altschul, et al. (1990) J. Mol. Biol. 215:403-410. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul, et al., *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, e.g., for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul (1993) Proc. Nat'l. Acad. Sci. USA 90:5873-5877). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between

two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. Log values may be negative
5 large numbers, e.g., 5, 10, 20, 30, 40, 40, 70, 90, 110, 150, 170, etc.

An indication that two nucleic acid sequences are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid. Thus, a polypeptide is typically substantially identical to a second polypeptide, e.g., where the two
10 peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequences.

15 A "host cell" is a naturally occurring cell or a transformed cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be cultured cells, explants, cells *in vivo*, and the like. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, and the like (see, e.g., the American Type Culture
20 Collection catalog or web site, www.atcc.org).

The terms "isolated," "purified," or "biologically pure" refer to material that is substantially or essentially free from components that normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid
25 chromatography. A protein or nucleic acid that is the predominant species present in a preparation is substantially purified. In particular, an isolated nucleic acid is separated from some open reading frames that naturally flank the gene and encode proteins other than protein encoded by the gene. The term "purified" in some embodiments denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Preferably, it means
30 that the nucleic acid or protein is at least about 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure. "Purify" or "purification" in other embodiments means removing at least one contaminant or component from the composition to be purified.

In this sense, purification does not require that the purified compound be homogeneous, e.g., 100% pure.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymer.

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function similarly to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refer to compounds that have the same basic chemical structure as a naturally occurring amino acid, e.g., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs may have modified R groups (e.g., norleucine) or modified peptide backbones, but retain some basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refer to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that function similarly to another amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical or associated, e.g., naturally contiguous, sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode most proteins. For instance, the codons GCA, GCC, GCG, and GCU each encode the amino acid alanine. Thus, at each position where an alanine is specified by a codon, the codon can be altered to another of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of

conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes silent variations of the nucleic acid. In certain contexts each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally similar molecule. Accordingly, a silent variation of a nucleic acid which encodes a polypeptide is implicit in a described sequence with respect to the expression product, but not necessarily with respect to actual probe sequences.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention. Typically conservative substitutions include for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (see, e.g., Creighton, Proteins (1984)).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, see, e.g., Alberts, et al. (1994) Molecular Biology of the Cell (3rd ed.) and Cantor and Schimmel (1980) Biophysical Chemistry Part I: The Conformation of Biological Macromolecules. "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that often form a compact unit of the polypeptide and are typically 25 to approximately 500 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of β -sheet and α -helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed, usually by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

"Nucleic acid" or "oligonucleotide" or "polynucleotide" or grammatical equivalents used herein means at least two nucleotides covalently linked together. Oligonucleotides are typically from about 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50 or more nucleotides in length, up to about 100 nucleotides in length. Nucleic acids and polynucleotides are polymers of any length, including longer lengths, e.g., 200, 300, 500, 1000, 2000, 3000, 5000, 7000, 10,000, etc. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs are included that may have at least one different linkage, e.g., phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages (see Eckstein (1992) Oligonucleotides and Analogues: A Practical Approach Oxford University Press); and peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, in Sanghui and Cook, eds. Carbohydrate Modifications in Antisense Research, ASC Symposium Series 580. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, e.g., to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature (T_m) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4° C drop in T_m for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9° C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. As will be appreciated by those in the art, the depiction of a single strand also defines the sequence of the complementary

strand; thus the sequences described herein also provide the complement of the sequence.

The nucleic acid may be DNA, both genomic and cDNA, RNA, or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine

5 hypoxanthine, isocytosine, isoguanine, etc. "Transcript" typically refers to a naturally occurring RNA, e.g., a pre-mRNA, hnRNA, or mRNA. As used herein, the term "nucleoside" includes nucleotides and nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occurring analog structures. Thus, e.g., the individual units of a peptide nucleic
10 acid, each containing a base, are referred to herein as a nucleoside.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, physiological, chemical, or other physical means. For example, useful labels include ^{32}P , fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins
15 or other entities which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide. The labels may be incorporated into the cancer nucleic acids, proteins, and antibodies. Many methods known in the art for conjugating the antibody to the label may be employed, including those methods described by Hunter, et al. (1962) Nature 144:945; David, et al. (1974) Biochemistry
20 13:1014-1021; Pain, et al. (1981) J. Immunol. Meth., 40:219-230; and Nygren (1982) J. Histochem. and Cytochem. 30:407-412.

An "effector" or "effector moiety" or "effector component" is a molecule that is bound (or linked, or conjugated), either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds, to an antibody.
25 The "effector" can be a variety of molecules including, e.g., detection moieties including radioactive compounds, fluorescent compounds, an enzyme or substrate, tags such as epitope tags, a toxin; activatable moieties, a chemotherapeutic agent; a lipase; an antibiotic; or a radioisotope emitting "hard" e.g., beta radiation.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either
30 covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe. Alternatively, method

using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin.

As used herein a "nucleic acid probe or oligonucleotide" is a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, e.g., through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, preferably one that does not functionally interfere with hybridization. Thus, e.g., probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. Probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled, e.g., with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled, e.g., with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence. Diagnosis or prognosis may be based at the genomic level, or at the level of RNA or protein expression.

The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, e.g., recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed *in vitro*, in general, by the manipulation of nucleic acid, e.g., using polymerases and endonucleases, in a form not normally found in nature. In this manner, operably linkage of different sequences is achieved. Thus an isolated nucleic acid, in a linear form, or an expression vector formed *in vitro* by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e., using the *in vivo* cellular machinery of the host cell rather than *in vitro* manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered

recombinant for the purposes of the invention. Similarly, a "recombinant protein" is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not normally found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences, e.g., from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein will often refer to two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

A "promoter" is typically an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, e.g., wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

An "expression vector" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed in operable linkage to a promoter.

The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule selectively to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to essentially no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in "Overview of principles of hybridization and the strategy of nucleic acid assays" in Tijssen (1993) Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes (vol. 24) Elsevier. Generally, stringent conditions are selected to be about 5-10° C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (e.g., 10 to 50 nucleotides) and at least about 60° C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is typically at least two times background, preferably 10 times background hybridization.

Exemplary stringent hybridization conditions are often: 50% formamide, 5x SSC, and 1% SDS, incubating at 42° C, or, 5x SSC, 1% SDS, incubating at 65° C, with wash in 0.2x SSC, and 0.1% SDS at 65° C. For PCR, a temperature of about 36° C is typical for low stringency amplification, although annealing temperatures may vary between about 32° C and 48° C depending on primer length. For high stringency PCR amplification, a temperature of about 62° C is typical, although high stringency annealing temperatures can range from about 50° C to about 65° C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90° C - 95° C for 0.5 - 2 min., an annealing phase lasting 0.5 - 2 min., and an extension phase of about 72° C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis, et al.(1990) PCR Protocols. A Guide to Methods and Applications.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This

occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37° C, and a wash in 1X SSC at 45° C. A positive hybridization is at least twice background. Alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., Ausubel, et al. (ed.) Current Protocols in Molecular Biology Lippincott.

The phrase "functional effects" in the context of assays for testing compounds that modulate activity of a lung cancer protein includes the determination of a parameter that is indirectly or directly under the influence of the lung cancer protein or nucleic acid, e.g., a physiological, enzymatic, functional, physical, or chemical effect, such as the ability to decrease lung cancer. It includes ligand binding activity; cell viability, cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein expression in cells undergoing metastasis, and other characteristics of lung cancer cells. "Functional effects" include *in vitro*, *in vivo*, and *ex vivo* activities.

By "determining the functional effect" is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a lung cancer protein sequence, e.g., physiological, functional, enzymatic, physical, or chemical effects. Such functional effects can be measured by many means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein, measuring inducible markers or transcriptional activation of the lung cancer protein; measuring binding activity or binding assays, e.g., binding to antibodies or other ligands, and measuring cellular proliferation. Determination of the functional effect of a compound on lung cancer can also be performed using lung cancer assays known to those of skill in the art such as an *in vitro* assays, e.g., cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein

expression in cells undergoing metastasis, and other characteristics of lung cancer cells. The functional effects can be evaluated by many means known to those skilled in the art, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, measurement of changes in RNA or protein levels for lung cancer-associated sequences, measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase, β -gal, GFP, and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

"Inhibitors", "activators", and "modulators" of lung cancer polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules or compounds identified using *in vitro* and *in vivo* assays of lung cancer polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of lung cancer proteins, e.g., antagonists. Antisense or inhibitory nucleic acids may seem to inhibit expression and subsequent function of the protein.

"Activators" are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate lung cancer protein activity. Inhibitors, activators, or modulators also include genetically modified versions of lung cancer proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, small chemical molecules and the like. Such assays for inhibitors and activators include, e.g., expressing the lung cancer protein *in vitro*, in cells, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above. Activators and inhibitors of lung cancer can also be identified by incubating lung cancer cells with the test compound and determining increases or decreases in the expression of 1 or more lung cancer proteins, e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 or more lung cancer proteins, such as lung cancer proteins encoded by the sequences set out in Tables 1A-16.

Samples or assays comprising lung cancer proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a polypeptide is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation of a lung cancer polypeptide is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more

preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

The phrase "changes in cell growth" refers to any change in cell growth and proliferation characteristics *in vitro* or *in vivo*, such as cell viability, formation of foci, anchorage independence, semi-solid or soft agar growth, changes in contact inhibition and density limitation of growth, loss of growth factor or serum requirements, changes in cell morphology, gaining or losing immortalization, gaining or losing tumor specific markers, ability to form or suppress tumors when injected into suitable animal hosts, and/or immortalization of the cell. See, e.g., Freshney (1994) Culture of Animal Cells a Manual of Basic Technique pp. 231-241 (3rd ed.).

"Tumor cell" refers to precancerous, cancerous, and normal cells in a tumor.

"Cancer cells," "transformed" cells, or "transformation" in tissue culture, refers to spontaneous or induced phenotypic changes that do not necessarily involve the uptake of new genetic material. Although transformation can arise from infection with a transforming virus and incorporation of new genomic DNA, or uptake of exogenous DNA, it can also arise spontaneously or following exposure to a carcinogen, thereby mutating an endogenous gene. Transformation is associated with phenotypic changes, such as immortalization of cells, aberrant growth control, nonmorphological changes, and/or malignancy (see, Freshney (1994) Culture of Animal Cells a Manual of Basic Technique (3rd ed.)).

"Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD, and IgE, respectively. Typically, the antigen-binding region of an antibody or its functional equivalent will be most critical in specificity and affinity of binding. See Paul, Fundamental Immunology.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible

for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, e.g., pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$, a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (see Paul (ed. 1999) Fundamental Immunology (4th ed.). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty, et al. (1990) Nature 348:552-554).

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (see, e.g., Kohler and Milstein (1975) Nature 256:495-497; Kozbor, et al. (1983) Immunology Today 4:72; Cole, et al. (1985), pp. 77-96 in Monoclonal Antibodies and Cancer Therapy; Coligan (1991 and supplements) Current Protocols in Immunology; Harlow and Lane (1988) Antibodies, A Laboratory Manual; and Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.)). Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty, et al. (1990) Nature 348:552-554; Marks, et al. (1992) Biotechnology 10:779-783).

A "chimeric antibody" is an antibody molecule in which, e.g., (a) the constant region, or a portion thereof, is altered, replaced, or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function, and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the

variable region, or a portion thereof, is altered, replaced, or exchanged with a variable region having a different or altered antigen specificity.

Identification of lung cancer-associated sequences

5 In one aspect, the expression levels of genes are determined in different patient samples for which diagnosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a "fingerprint" of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is
10 characteristic of the state of the cell. That is, normal tissue may be distinguished from cancerous or metastatic cancerous tissue, or metastatic cancerous tissue can be compared with tissue from surviving cancer patients. By comparing expression profiles of tissue in known different lung cancer states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained.
15 Molecular profiling may distinguish subtypes of a currently collective disease designation, e.g., different forms of lung cancer (chronic disease, adenocarcinoma, etc.)

The identification of sequences that are differentially expressed in lung cancer versus non-lung cancer tissue allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated: does a chemotherapeutic drug act to down-
20 regulate lung cancer, and thus tumor growth or recurrence, in a particular patient. Alternatively, a treatment step may induce other markers which may be used as targets to destroy tumor cells. Similarly, diagnosis and treatment outcomes may be done or confirmed by comparing patient samples with the known expression profiles. Malignant disease may be compared to non-malignant conditions. Metastatic tissue can also be analyzed to determine
25 the stage of lung cancer in the tissue, or origin of primary tumor, e.g., metastasis from a remote primary site. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to mimicking or altering a particular expression profile; e.g., screening can be done for drugs that suppress the lung cancer expression profile. This may be done by making biochips comprising sets of the important lung cancer genes,
30 which can then be used in these screens. PCR methods may be applied with selected primer pairs, and analysis may be of RNA or of genomic sequences. These methods can also be done on the protein basis; that is, protein expression levels of the lung cancer proteins can be evaluated for diagnostic purposes or to screen candidate agents. In addition, the lung cancer

nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the lung cancer proteins (including antibodies and other modulators thereof) administered as therapeutic drugs or as protein or DNA vaccines.

5 Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in lung cancer relative to normal tissues and/or non-malignant lung disease, or in different types of lung disease, herein termed "lung cancer sequences." As outlined below, lung cancer sequences include those that are up-regulated (i.e., expressed at a higher level) in lung cancer, as well as those that are down-regulated (i.e., expressed at a lower level). In a preferred embodiment, the lung cancer sequences are from humans; 10 however, as will be appreciated by those in the art, lung cancer sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other lung cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc.) and pets (dogs, cats, etc.). Lung cancer sequences from other organisms may be 15 obtained using the techniques outlined below.

Lung cancer sequences can include both nucleic acid and amino acid sequences. As will be appreciated by those in the art and is more fully outlined below, lung cancer nucleic acid sequences are useful in a variety of applications, including diagnostic applications, 20 which will detect naturally occurring nucleic acids, as well as screening applications; e.g., biochips comprising nucleic acid probes or PCR microtiter plates with selected probes to the lung cancer sequences can be generated.

A lung cancer sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the lung cancer sequences outlined herein. Such 25 homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, e.g., using homology programs or hybridization conditions.

For identifying lung cancer-associated sequences, the lung cancer screen typically includes comparing genes identified in different tissues, e.g., normal and cancerous tissues, 30 cancer and non-malignant conditions, non-malignant conditions and normal tissues, or tumor tissue samples from patients who have metastatic disease vs. non metastatic tissue. Other suitable tissue comparisons include comparing lung cancer samples with metastatic cancer samples from other cancers, such as, breast, other gastrointestinal cancers, prostate, ovarian,

etc. Samples of, non metastatic disease tissue and tissue undergoing metastasis are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as is known in the art for the preparation of mRNA. Suitable biochips are commercially available, e.g., from Affymetrix, Santa Clara, CA. Gene expression profiles as described herein are generated and the data analyzed.

In one embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, preferably normal lung, but also including, and not limited to colon, heart, brain, liver, breast, kidney, muscle, prostate, small intestine, large intestine, spleen, bone, and/or placenta. In a preferred embodiment, those genes identified during the lung cancer screen that are expressed in significant amounts in other tissues (e.g., essential organs) are removed from the profile, although in some embodiments, this is not necessary (e.g., where organs may be dispensable at a later stage of life). That is, when screening for drugs, it is usually preferable that the target expression be disease specific, to minimize possible side effects on other organs.

In a preferred embodiment, lung cancer sequences are those that are up-regulated in lung cancer; that is, the expression of these genes is higher in cancerous tissue than in normal lung or other tissue. "Up-regulation" as used herein means, when the ratio is presented as a number greater than one, that the ratio is greater than one, preferably 1.5 or greater, more preferably 2.0 or greater. Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal. Unigene cluster identification numbers and accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is known in the art, see, e.g., Benson, DA, et al (1998) *Nucleic Acids Research* 26:1-7 and <http://www.ncbi.nlm.nih.gov/>. Sequences are also available in other databases, e.g., European Molecular Biology Laboratory (EMBL) and DNA Database of Japan (DDBJ). Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal. In some situations, the sequences may be derived from assembly of available sequences or be predicted from genomic DNA using exon prediction algorithms, such as FGENESH (Salamov and Solovyev (2000) *Genome Res.* 10:516-522). In other situations, sequences have been derived from cloning and sequencing of isolated nucleic acids.

In another preferred embodiment, lung cancer sequences are those that are down-regulated in the lung cancer; that is, the expression of these genes is lower in cancerous tissue

or normal lung or other tissue. "Down-regulation" as used herein means, when the ratio is presented as a number greater than one, that the ratio is greater than one, preferably 1.5 or greater, more preferably 2.0 or greater, or, when the ratio is presented as a number less than one, that the ratio is less than one, preferably 0.5 or less, more preferably 0.25 or less.

5

Informatics

The ability to identify genes that are over or under expressed in lung cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure, biosensor development, and other related areas. For example, the expression profiles can be used in diagnostic or prognostic evaluation of patients with lung cancer. Or as another example, subcellular toxicological information can be generated to better direct drug structure and activity correlation (see Anderson (1998) Pharmaceutical Proteomics: Targets, Mechanism, and Function, paper presented at the IBC Proteomics conference, Coronado, CA (June 11-12, 1998)). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical exposures and likely tolerable exposure thresholds (see U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (e.g., nucleic acids, saccharides, lipids, drugs, and the like).

Thus, in another embodiment, the present invention provides a database that includes at least one set of assay data. The data contained in the database is acquired, e.g., using array analysis either singly or in a library format. The database can be in a form in which data can be maintained and transmitted, but is preferably an electronic database. The electronic database of the invention can be maintained on any electronic device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

The focus of the present section on databases that include peptide sequence data is for clarity of illustration only. It will be apparent to those of skill in the art that similar databases can be assembled for assay data acquired using an assay of the invention.

The compositions and methods for identifying and/or quantitating the relative and/or absolute abundance of a variety of molecular and macromolecular species from a biological sample representing lung cancer, i.e., the identification of lung cancer-associated sequences described herein, provide an abundance of information, which can be correlated with

pathological conditions, predisposition to disease, drug testing, therapeutic monitoring, gene-disease causal linkages, identification of correlates of immunity and physiological status, among others. Although the data generated from the assays of the invention is suited for manual review and analysis, in a preferred embodiment, data processing using high-speed computers is utilized.

An array of methods for indexing and retrieving biomolecular information is known in the art. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more sequencing projects for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for making a retrieval of a gene sequence similar to a sequence data item in a gene database based on the degree of similarity between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent 5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension. U.S. Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which can be viewed as a tree structure or as the merger of two or more such tree structures.

See also Mount, et al. (2001) Bioinformatics; Durbin, et al. (eds., 1999) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids (; Baxeavanis and Oeullette (eds., 1998) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins); Rashidi and Buehler (1999) Bioinformatics: Basic Applications in Biological Science and Medicine; Setubal, et al. (eds 1997) Introduction to Computational Molecular Biology; Misener and Krawetz (eds, 2000) Bioinformatics: Methods and Protocols; Higgins and Taylor (eds., 2000) Bioinformatics: Sequence, Structure, and Databanks: A Practical

Approach; Brown (2001) Bioinformatics: A Biologist's Guide to Biocomputing and the Internet; Han and Kamber (2000) Data Mining: Concepts and Techniques (2000); and Waterman (1995) Introduction to Computational Biology: Maps, Sequences, and Genomes.

The present invention provides a computer database comprising a computer and
5 software for storing in computer-retrievable form assay data records cross-tabulated, e.g., with data specifying the source of the target-containing sample from which each sequence specificity record was obtained.

In an exemplary embodiment, at least one of the sources of target-containing sample is from a control tissue sample known to be free of pathological disorders. In a variation, at
10 least one of the sources is a known pathological tissue specimen, e.g., a neoplastic lesion or another tissue specimen to be analyzed for lung cancer. In another variation, the assay records cross-tabulate one or more of the following parameters for each target species in a sample: (1) a unique identification code, which can include, e.g., a target molecular structure and/or characteristic separation coordinate (e.g., electrophoretic coordinates); (2) sample
15 source; and (3) absolute and/or relative quantity of the target species present in the sample.

The invention also provides for the storage and retrieval of a collection of target data in a computer data storage apparatus, which can include magnetic disks, optical disks, magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic bubble memory devices, and other data storage devices, including CPU registers and on-CPU
20 data storage arrays. Typically, the target data records are stored as a bit pattern in an array of magnetic domains on a magnetizable medium or as an array of charge states or transistor gate states, such as an array of cells in a DRAM device (e.g., each cell comprised of a transistor and a charge storage area, which may be on the transistor). In one embodiment, the invention provides such storage devices, and computer systems built therewith, comprising a bit pattern
25 encoding a protein expression fingerprint record comprising unique identifiers for at least 10 target data records cross-tabulated with target source.

When the target is a peptide or nucleic acid, the invention preferably provides a method for identifying related peptide or nucleic acid sequences, comprising performing a computerized comparison between a peptide or nucleic acid sequence assay record stored in
30 or retrieved from a computer storage device or database and at least one other sequence. The comparison can include a sequence analysis or comparison algorithm or computer program embodiment thereof (e.g., FASTA, TFASTA, GAP, BESTFIT) and/or the comparison may

be of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

The invention also preferably provides a magnetic disk, such as an IBM-compatible (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format (e.g., Linux, SunOS, Solaris, AIX, SCO Unix, VMS, MV, Macintosh, etc.) floppy diskette or hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis, comparison, or relative quantitation method.

The invention also provides a network, comprising a plurality of computing devices linked via a data link, such as an Ethernet cable (coax or 10BaseT), telephone line, ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (e.g., computer, disk array, etc.) comprises a pattern of magnetic domains (e.g., magnetic disk) and/or charge domains (e.g., an array of DRAM cells) composing a bit pattern encoding data acquired from an assay of the invention.

The invention also provides a method for transmitting assay data that includes generating an electronic signal on an electronic communications device, such as a modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal includes (in native or encrypted format) a bit pattern encoding data from an assay or a database comprising a plurality of assay results obtained by the method of the invention.

In a preferred embodiment, the invention provides a computer system for comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the degree of identity and gap weight to the target data. A central processor is preferably initialized to load and execute the computer program for alignment and/or comparison of the assay results.

Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

The target data or record and the computer program can be transferred to secondary memory, which is typically random access memory (e.g., DRAM, SRAM, SGRAM, or SDRAM). Targets are ranked according to the degree of correspondence between a selected assay characteristic (e.g., binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor can be a conventional computer (e.g., Intel Pentium, PowerPC, Alpha, PA-8000, SPARC,

MIPS 4400, MIPS 10000, VAX, etc.); a program can be a commercial or public domain molecular biology software package (e.g., UWGCG Sequence Analysis Software, Darwin); a data file can be an optical or magnetic disk, a data server, a memory device (e.g., DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, etc.); an I/O device can be a terminal comprising a video display and a keyboard, a modem, an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which may be stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values.

15 Characteristics of lung cancer-associated proteins

Lung cancer proteins of the present invention may be classified as secreted proteins, transmembrane proteins or intracellular proteins. In one embodiment, the lung cancer protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, e.g., signaling pathways); aberrant expression of such proteins often results in unregulated or dysregulated cellular processes (see, e.g., Alberts (ed. 1994) Molecular Biology of the Cell (3d ed.)). For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterizing proteins is the presence in the proteins of one or more structural motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2

domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by one of ordinary skill in the art, these motifs can be identified on the basis of amino acid sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate. One useful database is Pfam (protein families), which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Versions are available via the internet from Washington University in St. Louis, the Sanger Center in England, and the Karolinska Institute in Sweden (see, e.g., Bateman, et al (2000) *Nuc. Acids Res.* 28:263-266; Sonnhammer, et al. (1997) *Proteins* 28:405-420; Bateman, et al. (1999) *Nuc. Acids Res.* 27:260-262; and Sonnhammer, et al. (1998) *Nuc. Acids Res.* 26:320-322).

In another embodiment, the lung cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels, pumps, and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Characteristics of transmembrane domains include approximately 17 consecutive hydrophobic amino acids that may be followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the

localization and number of transmembrane domains within the protein may be predicted (see, e.g., PSORT web site <http://psort.nibb.ac.jp/>).

The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF, and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, hormones, neurotrophic factors and the like. Extracellular domains also bind to cell-associated molecules. In this respect, they may mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell, e.g., via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains may also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

Lung cancer proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for extracellular immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins can be also useful in imaging modalities. Antibodies may be used to label such readily accessible proteins *in situ* or in histological analysis. Alternatively, antibodies can also label intracellular proteins, in which case analytical samples are typically permeabilized to provide access to intracellular proteins. In addition, some membrane proteins can be processed to release a soluble protein, or to expose a residual fragment. Released soluble proteins may be useful diagnostic markers, processed residual protein fragments may be useful lung markers of disease.

It will also be appreciated by those in the art that a transmembrane protein can be made soluble by removing transmembrane sequences, e.g., through recombinant methods. Furthermore, transmembrane proteins that have been made soluble can be made to be secreted through recombinant means by adding an appropriate signal sequence.

In another embodiment, the lung cancer proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins may have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; e.g., if circulating, they often serve to transmit

signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor), an endocrine manner (acting on cells at a distance, e.g., secretion into the blood stream), or exocrine (secretion, e.g., through a duct or to adjacent epithelial surface as sweat glands, sebaceous glands, pancreatic ducts, lacrimal glands, mammary glands, salivary producing glands of the ear, etc.). Thus secreted molecules often find use in modulating or altering numerous aspects of physiology. Lung cancer proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, e.g., for blood, plasma, serum, or stool tests. Those which are enzymes may be antibody or small molecule targets. Others may be useful as vaccine targets, e.g., via CTL mechanisms.

Use of lung cancer nucleic acids

As described above, lung cancer sequence is initially identified by substantial nucleic acid and/or amino acid sequence homology or linkage to the lung cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions. Typically, linked sequences on a mRNA are found on the same molecule.

The lung cancer nucleic acid sequences of the invention, e.g., the sequences in Tables 1A-16, can be fragments of larger genes, i.e., they are nucleic acid segments. "Genes" in this context includes coding regions, non-coding regions, and mixtures of coding and non-coding regions. Accordingly, as will be appreciated by those in the art, using the sequences provided herein, extended sequences, in either direction, of the lung cancer genes can be obtained, using techniques well known in the art for cloning either longer sequences or the full length sequences; see Ausubel, et al., *supra*. Much can be done by informatics and many sequences can be clustered to include multiple sequences corresponding to a single gene, e.g., systems such as UniGene (see, <http://www.ncbi.nlm.nih.gov/UniGene/>).

Once a lung cancer nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire lung cancer nucleic acid coding regions or the entire mRNA sequence. Once isolated from its natural source, e.g., contained within a plasmid or other vector or excised therefrom as a linear nucleic acid segment, the recombinant lung cancer nucleic acid can be further-used as a probe to identify and isolate

other lung cancer nucleic acids, e.g., extended coding regions. It can also be used as a "precursor" nucleic acid to make modified or variant lung cancer nucleic acids and proteins.

The lung cancer nucleic acids of the present invention are used in several ways. In a first embodiment, nucleic acid probes to the lung cancer nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, e.g., for gene therapy, RNAi, vaccine, and/or antisense applications. Alternatively, the lung cancer nucleic acids that include coding regions of lung cancer proteins can be put into expression vectors for the expression of lung cancer proteins, again for screening purposes or for administration to a patient.

In a preferred embodiment, nucleic acid probes to lung cancer nucleic acids (both the nucleic acid sequences outlined in the figures and/or the complements thereof) are made. The nucleic acid probes attached to the biochip are designed to be substantially complementary to the lung cancer nucleic acids, i.e., the target sequence (either the target sequence of the sample or to other probe sequences, e.g., in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary" herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under appropriate reaction conditions, particularly high stringency conditions, as outlined herein.

A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8 to about 100 bases long, with from about 10 to about 80 bases being preferred, and from about 30 to about 50 bases being particularly preferred. That is, generally complements of ORFs or whole genes are not used. In some embodiments, nucleic acids of lengths up to hundreds of bases can be used.

In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a

particular target. The probes can be overlapping (i.e., have some sequence in common), or separate. In some cases, PCR primers may be used to amplify signal for higher sensitivity.

As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By "immobilized" and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can typically be covalent or non-covalent. By "non-covalent binding" and grammatical equivalents herein is typically meant one or more of electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By "covalent binding" and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

In general, the probes are attached to a biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

The biochip comprises a suitable solid substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant a material that can be modified for the attachment or association of the nucleic acid probes and is amenable to at least one detection method. Often the substrate may contain discrete individual sites appropriate for individual partitioning and identification. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. A preferred substrate is described in US application entitled Reusable Low Fluorescent Plastic Biochip, U.S.

Application Serial No. 09/270,214, filed March 15, 1999, herein incorporated by reference in its entirety.

Generally the substrate is planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, e.g., the biochip is derivatized with a chemical functional group including, but not limited to, amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, e.g., using linkers as are known in the art; e.g., homo-or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

In this embodiment, oligonucleotides are synthesized, and then attached to the surface of the solid support. Either the 5' or 3' terminus may be attached to the solid support, or attachment may be via linkage to an internal nucleoside.

In another embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized *in situ*, using known photolithographic techniques, such as those described in WO 95/25116; WO 95/35505; U.S. Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affymetrix GeneChip™ technology.

Often, amplification-based assays are performed to measure the expression level of lung cancer-associated sequences. These assays are typically performed in conjunction with

reverse transcription. In such assays, a lung cancer-associated nucleic acid sequence acts as a template in an amplification reaction (e.g., Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the amount of lung cancer-associated RNA. Methods of quantitative amplification are well known to those of skill in the art. Detailed protocols for quantitative PCR are provided, e.g., in Innis, et al. (1990) PCR Protocols, A Guide to Methods and Applications.

In some embodiments, a TaqMan based assay is used to measure expression. TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, e.g., AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (see, e.g., literature provided by Perkin-Elmer, e.g., www2.perkin-elmer.com).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (see Wu and Wallace (1989) Genomics 4:560, Landegren, et al. (1988) Science 241:1077, and Barringer, et al. (1990) Gene 89:117), transcription amplification (Kwoh, et al. (1989) Proc. Natl. Acad. Sci. USA 86:1173), self-sustained sequence replication (Guatelli, et al. (1990) Proc. Nat. Acad. Sci. USA 87:1874), dot PCR, and linker adapter PCR, etc.

Expression of lung cancer proteins from nucleic acids

In a preferred embodiment, lung cancer nucleic acids, e.g., encoding lung cancer proteins, are used to make a variety of expression vectors to express lung cancer proteins which can then be used in screening assays, as described below. Expression vectors and recombinant DNA technology are well known to those of skill in the art (see, e.g., Ausubel, *supra*, and Fernandez and Hoeffler (eds 1999) Gene Expression Systems) and are used to express proteins. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the lung cancer protein. The term "control sequences" refers to DNA

sequences used for the expression of an operably linked coding sequence in a particular host organism. Control sequences that are suitable for prokaryotes, e.g., include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

5 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding
10 sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is typically accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in
15 accordance with conventional practice. Transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the lung cancer protein. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

In general, transcriptional and translational regulatory sequences may include, but are
20 not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

Promoter sequences may be either constitutive or inducible promoters. The promoters
25 may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

In addition, an expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two
30 organisms, e.g., in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector often contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct. The integrating

vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art (e.g., Fernandez and Hoeffler, *supra*).

5 In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The lung cancer proteins of the present invention are usually produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a lung cancer protein, under the appropriate conditions to induce or cause expression of the lung cancer protein. Conditions appropriate for lung cancer protein expression will vary with the
10 choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation or optimization. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate
15 growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

Appropriate host cells include yeast, bacteria, archaeobacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Saccharomyces cerevisiae*
20 and other yeasts, *E. coli*, *Bacillus subtilis*, S9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, HUVEC (human umbilical vein endothelial cells), THP1 cells (a macrophage cell line) and various other human cells and cell lines.

In a preferred embodiment, the lung cancer proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral and
25 adenoviral systems. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter (see, e.g., Fernandez and Hoeffler, *supra*). Typically, transcription termination and
30 polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenylation signals include those derived from SV40.

The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, lung cancer proteins are expressed in bacterial systems. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; e.g., the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the lung cancer protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others (e.g., Fernandez and Hoeffler, *supra*). The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

In one embodiment, lung cancer proteins are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

In a preferred embodiment, lung cancer protein is produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*,

Kluyveromyces fragilis and *K. lactis*, *Pichia guillermondii*, and *P. pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

The lung cancer protein may also be made as a fusion protein, using techniques well known in the art. Thus, e.g., for the creation of monoclonal antibodies, if the desired epitope is small, the lung cancer protein may be fused to a carrier protein to form an immunogen. Alternatively, the lung cancer protein may be made as a fusion protein to increase expression for affinity purification purposes, or for other reasons. For example, when the lung cancer protein is a lung cancer peptide, the nucleic acid encoding the peptide may be linked to other nucleic acid for expression purposes.

In a preferred embodiment, the lung cancer protein is purified or isolated after expression. Lung cancer proteins may be isolated or purified in a variety of appropriate ways. Standard purification methods include electrophoretic, molecular, immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the lung cancer protein may be purified using a standard anti-lung cancer protein antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. For general guidance in suitable purification techniques, see Scopes (1982) Protein Purification. The degree of purification necessary will vary depending on the use of the lung cancer protein. In some instances no purification will be necessary.

Once expressed and purified if necessary, the lung cancer proteins and nucleic acids are useful in a number of applications. They may be used as immunoselection reagents, as vaccine reagents, as screening agents, therapeutic entities, for production of antibodies, as transcription or translation inhibitors, etc.

Variants of lung cancer proteins

In one embodiment, the lung cancer proteins are derivative or variant lung cancer proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative lung cancer peptide will often contain at least one amino acid substitution, deletion or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion or deletion may occur at a particular residue within the lung cancer peptide.

Also included within one embodiment of lung cancer proteins of the present invention are amino acid sequence variants. These variants typically fall into one or more of three

classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the lung cancer protein, using cassette or PCR mutagenesis or other techniques, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above.

5 However, variant lung cancer protein fragments having up to about 100-150 residues may be prepared by *in vitro* synthesis. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the lung cancer protein amino acid sequence. The variants typically exhibit a similar qualitative biological activity as the naturally occurring analogue,

10 although variants can also be selected which have modified characteristics as will be more fully outlined below.

While the site or region for introducing an amino acid sequence variation is often predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be

15 conducted at the target codon or region and the expressed lung cancer variants screened for the optimal combination of desired activity. Techniques exist for making substitution mutations at predetermined sites in DNA having a known sequence, e.g., M13 primer mutagenesis and PCR mutagenesis. Screening of mutants is often done using assays of lung cancer protein activities.

20 Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger insertions may be occasionally tolerated. Deletions generally range from about 1 to about 20 residues, although in some cases deletions may be much larger.

Substitutions, deletions, insertions or any combination thereof may be used to arrive

25 at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. Larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of a lung cancer protein are desired, substitutions are generally made in accordance with the amino acid substitution chart provided in the definition section.

30 Variants typically exhibit essentially the same qualitative biological activity and will elicit the same immune response as a naturally-occurring analog, although variants also are selected to modify the characteristics of lung cancer proteins as needed. Alternatively, the

variant may be designed or reorganized such that a biological activity of the lung cancer protein is altered. For example, glycosylation sites may be added, altered, or removed.

Covalent modifications of lung cancer polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a lung cancer polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a lung cancer polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking lung cancer polypeptides to a water-insoluble support matrix or surface for use in a method for purifying anti-lung cancer polypeptide antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, e.g., esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-((p-azidophenyl)dithio)propioimide.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of serinyl, threonyl or tyrosyl residues, methylation of the γ -amino groups of lysine, arginine, and histidine side chains (Creighton (1983) Proteins: Structure and Molecular Properties, pp. 79-86), acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the lung cancer polypeptide encompassed by this invention is an altered native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended herein to mean adding to or deleting one or more carbohydrate moieties of a native sequence lung cancer polypeptide. Glycosylation patterns can be altered in many ways. For example the use of different cell types to express lung cancer-associated sequences can result in different glycosylation patterns.

Addition of glycosylation sites to lung cancer polypeptides may also be accomplished by altering the amino acid sequence thereof. The alteration may be made, e.g., by the addition of, or substitution by, one or more serine or threonine residues to the native sequence lung cancer polypeptide (for O-linked glycosylation sites). The lung cancer amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the lung cancer polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the lung cancer polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330, and in Aplin and Wriston (1981) CRC Crit. Rev. Biochem., pp. 259-306.

- 5 Removal of carbohydrate moieties present on the lung cancer polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al. (1987) Arch. Biochem. Biophys., 259:52 and by Edge, et al. (1981) Anal. Biochem., 118:131.
- 10 Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura, et al. (1987) Meth. Enzymol., 138:350.

- Another type of covalent modification of lung cancer comprises linking the lung cancer polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent
- 15 Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192, or 4,179,337.

- Lung cancer polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a lung cancer polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric
- 20 molecule comprises a fusion of a lung cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the lung cancer polypeptide. The presence of such epitope-tagged forms of a lung cancer polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the lung
- 25 cancer polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a lung cancer polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

- 30 Various tag polypeptides and their respective antibodies are well known and examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; HIS6 and metal chelation tags, the flu HA tag polypeptide and its antibody 12CA5 (Field, et al. (1988) Mol. Cell. Biol. 8:2159-2165); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies

thereto (Evan, et al. (1985) Molecular and Cellular Biology 5:3610-3616); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky, et al. (1990) Protein Engineering 3(6):547-553). Other tag polypeptides include the Flag-peptide (Hopp, et al. (1988) BioTechnology 6:1204-1210); the KT3 epitope peptide (Martin, et al. (1992) Science 255:192-194); tubulin epitope peptide (Skinner, et al. (1991) J. Biol. Chem. 266:15163-15166); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth, et al. (1990) Proc. Nat'l Acad. Sci. USA 87:6393-6397).

Also included are other lung cancer proteins of the lung cancer family, and lung cancer proteins from other organisms, which are cloned and expressed as outlined below.

Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related lung cancer proteins from primates or other organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include unique areas of the lung cancer nucleic acid sequence. As is generally known in the art, preferred PCR primers are from about 15 to about 35 nucleotides in length, with from about 20 to about 30 being preferred, and may contain inosine as needed. PCR reaction conditions are well known in the art (e.g., Innis, PCR Protocols, *supra*).

Antibodies to lung cancer proteins

In a preferred embodiment, when a lung cancer protein is to be used to generate antibodies, e.g., for immunotherapy or immunodiagnosis, the lung cancer protein should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is typically meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies made to a smaller lung cancer protein will be able to bind to the full-length protein, particularly linear epitopes. In a preferred embodiment, the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity.

Methods of preparing polyclonal antibodies are well known (e.g., Coligan, *supra*; and Harlow and Lane, *supra*). Polyclonal antibodies can be raised in a mammal, e.g., by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of Tables 1A-16 or fragment thereof or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal

being immunized. Immunogenic proteins include, e.g., keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Adjuvants include, e.g., Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art.

The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein (1975) Nature 256:495. In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*. The immunizing agent will typically include a polypeptide encoded by a nucleic acid of the tables, or fragment thereof, or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding (1986) Monoclonal Antibodies: Principles and Practice, pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovin, or primate origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

In one embodiment, the antibodies are bispecific antibodies. Bispecific antibodies are typically monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens or that have binding specificities for two epitopes on the same antigen. In one embodiment, one of the binding specificities is for a protein encoded by a nucleic acid of the tables or a fragment thereof, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific. Alternatively, tetramer-type technology may create multivalent reagents.

In a preferred embodiment, the antibodies to lung cancer protein are capable of reducing or eliminating a biological function of a lung cancer protein, in a naked form or conjugated to an effector moiety. That is, the addition of anti-lung cancer protein antibodies (either polyclonal or preferably monoclonal) to lung cancer tissue (or cells containing lung cancer) may reduce or eliminate the lung cancer. Generally, at least a 25% decrease in activity, growth, size or the like is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

In a preferred embodiment the antibodies to the lung cancer proteins are humanized antibodies (e.g., Xenerex Biosciences, Medarex, Inc., Abgenix, Inc., Protein Design Labs, Inc.) Humanized forms of non-human (e.g., murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of a human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, a humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework (FR) regions are those of a human immunoglobulin consensus sequence. A humanized antibody optimally also will typically comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-329; and Presta (1992) Curr. Op. Struct. Biol. 2:593-596). Humanization can be performed following the method of Winter and co-workers (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-327; Verhoeven, et al. (1988) Science 239:1534-1536), by substituting rodent CDRs or CDR sequences for corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by corresponding sequence from a non-human species.

Human-like antibodies can also be produced using various techniques known in the art, including phage display libraries (Hoogenboom and Winter (1991) J. Mol. Biol. 227:381; Marks, et al. (1991) J. Mol. Biol. 222:581). The techniques of Cole, et al. and Boerner, et al. are also available for the preparation of human monoclonal antibodies (Cole, et al. (1985) Monoclonal Antibodies and Cancer Therapy, p. 77 and Boerner, et al. (1991) J. Immunol. 147(1):86-95). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in nearly all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, e.g., in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks, et al. (1992) Bio/Technology 10:779-783; Lonberg, et al. (1994) Nature 368:856-859; Morrison (1994) Nature 368:812-13; Fishwild, et al. (1996) Nature Biotechnology 14:845-51; Neuberger (1996) Nature Biotechnology 14:826; and Lonberg and Huszar (1995) Intern. Rev. Immunol. 13:65-93.

By immunotherapy is meant treatment of lung cancer with an antibody raised against a lung cancer proteins. As used herein, immunotherapy can be passive or active. Passive immunotherapy as defined herein is the passive transfer of antibody to a recipient (patient). Active immunization is the induction of antibody and/or T-cell responses in a recipient (patient). Induction of an immune response is the result of providing the recipient with an antigen to which antibodies are raised. The antigen may be provided by injecting a polypeptide against which antibodies are desired to be raised into a recipient, or contacting the recipient with a nucleic acid capable of expressing the antigen and under conditions for expression of the antigen, leading to an immune response.

In a preferred embodiment the lung cancer proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory, antibodies used for treatment, may bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted lung cancer protein.

In another preferred embodiment, the lung cancer protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment may bind the extracellular domain of the lung cancer protein and prevent it from binding to other proteins, such as circulating ligands or cell-associated molecules. The

antibody may cause down-regulation of the transmembrane lung cancer protein. The antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the lung cancer protein. The antibody may be an antagonist of the lung cancer protein or may prevent activation of a transmembrane lung cancer protein, or may induce or suppress a particular cellular pathway. In some embodiments, when the antibody prevents the binding of other molecules to the lung cancer protein, the antibody prevents growth of the cell. The antibody may also be used to target or sensitize the cell to cytotoxic agents; including, but not limited to TNF- α , TNF- β , IL-1, INF- γ , and IL-2, or chemotherapeutic agents including 5FU, vinblastine, actinomycin D, cisplatin, methotrexate, and the like. In some instances the antibody may belong to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity or antigen-dependent cytotoxicity (ADCC). Thus, lung cancer may be treated by administering to a patient antibodies directed against the transmembrane lung cancer protein. Antibody-labeling may activate a co-toxin, localize a toxin payload, or otherwise provide means to locally ablate cells.

In another preferred embodiment, the antibody is conjugated to an effector moiety. The effector moiety can be various molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the therapeutic moiety is a small molecule that modulates the activity of a lung cancer protein. In another aspect the therapeutic moiety may modulate an activity of molecules associated with or in close proximity to a lung cancer protein. The therapeutic moiety may inhibit enzymatic or signaling activity such as protease or collagenase activity associated with lung cancer.

In a preferred embodiment, the therapeutic moiety can also be a cytotoxic agent. In this method, targeting the cytotoxic agent to lung cancer tissue or cells results in a reduction in the number of afflicted cells, thereby reducing symptoms associated with lung cancer. Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their corresponding fragments include diphtheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, crotin, phenomycin, enomycin, saporin, auristatin, and the like. Cytotoxic agents also include radiochemicals made by conjugating radioisotopes to antibodies raised against lung cancer proteins, or binding of a radionuclide to a chelating agent that has been covalently attached to the antibody. Targeting the therapeutic moiety to transmembrane lung cancer proteins not only serves to increase the local concentration of therapeutic moiety in the lung cancer

afflicted area, but also serves to reduce deleterious side effects that may be associated with the untargeted therapeutic moiety.

In another preferred embodiment, the lung cancer protein against which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated to a protein or other entity which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the lung cancer protein can be targeted within a cell, i.e., the nucleus, an antibody theretomay contain a signal for that target localization, i.e., a nuclear localization signal.

The lung cancer antibodies of the invention specifically bind to lung cancer proteins. By "specifically bind" herein is meant that the antibodies bind to the protein with a K_d of at least about 0.1 mM, more usually at least about 1 μ M, preferably at least about 0.1 μ M or better, and most preferably, 0.01 μ M or better. Selectivity of binding to the specific target and not to related other sequences is also important.

Detection of lung cancer sequence for diagnostic and therapeutic applications

In one aspect, the RNA expression levels of genes are determined for different cellular states in the lung cancer phenotype. Expression levels of genes in normal tissue (e.g., not undergoing lung cancer), in lung cancer tissue (and in some cases, for varying severities of lung cancer that relate to prognosis, as outlined below), or in non-malignant disease are evaluated to provide expression profiles. A gene expression profile of a particular cell state or point of development is essentially a "fingerprint" of the state of the cell. While two states may have a particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be performed or confirmed to determine whether a tissue sample has the gene expression profile of normal or cancerous tissue. This will provide for molecular diagnosis of related conditions.

"Differential expression," or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, e.g.,

normal versus lung cancer tissue. Genes may be turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both. Alternatively, the difference in expression may be quantitative, e.g., in that expression is increased or decreased; i.e., gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays, Lockhart (1996) Nature Biotechnology 14:1675-1680, hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, northern analysis and RNase protection. As outlined above, preferably the change in expression (i.e., upregulation or downregulation) is typically at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation may be at the gene transcript or the protein level. The amount of gene expression may be monitored using nucleic acid probes to the RNA or DNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, e.g., with antibodies to the lung cancer protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to lung cancer genes, e.g., those identified as being important in a lung cancer or disease phenotype, can be evaluated in a lung cancer diagnostic test. In a preferred embodiment, gene expression monitoring is performed simultaneously on a number of genes.

The lung cancer nucleic acid probes may be attached to biochips as outlined herein for the detection and quantification of lung cancer sequences in a particular cell. The assays are further described below in the example. PCR techniques can be used to provide greater sensitivity. Multiple protein expression monitoring can be performed as well. Similarly, these assays may be performed on an individual basis as well.

In a preferred embodiment nucleic acids encoding the lung cancer protein are detected. Although DNA or RNA encoding the lung cancer protein may be detected, of particular interest are methods wherein an mRNA encoding a lung cancer protein is detected.

Probes to detect mRNA can be a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method detection of the mRNA is performed *in situ*. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a lung cancer protein is detected by binding the digoxigenin with an anti-digoxigenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) are used in diagnostic assays. The lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in diagnostic assays. This can be performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, lung cancer proteins, including intracellular, transmembrane, or secreted proteins, find use as markers of lung cancer, e.g., for prognostic or diagnostic purposes. Detection of these proteins in putative lung cancer tissue allows for detection, prognosis, or diagnosis of lung cancer or similar disease, and perhaps for selection of therapeutic strategy. In one embodiment, antibodies are used to detect lung cancer proteins. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be another type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the lung cancer protein is detected, e.g., by immunoblotting with antibodies raised against the lung cancer protein. Methods of immunoblotting are well known to those of ordinary skill in the art.

In another preferred method, antibodies to the lung cancer protein find use in *in situ* imaging techniques, e.g., in histology (e.g., Asai (ed. 1993) Methods in Cell Biology:

Antibodies in Cell Biology, volume 37. In this method cells are contacted with from one to many antibodies to the lung cancer protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label, e.g., multicolor fluorescence or confocal imaging. In another method the primary antibody to the lung cancer protein(s) contains a detectable label, e.g., an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of lung cancer proteins. Many other histological imaging techniques are also provided by the invention.

In a preferred embodiment the label is detected in a fluorometer which has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

In another preferred embodiment, antibodies find use in diagnosing lung cancer from blood, serum, plasma, stool, and other samples. Such samples, therefore, are useful as samples to be probed or tested for the presence of lung cancer proteins. Antibodies can be used to detect a lung cancer protein by previously described immunoassay techniques including ELISA, immunoblotting (western blotting), immunoprecipitation, BIACORE technology and the like. Conversely, the presence of antibodies may indicate an immune response against an endogenous lung cancer protein or vaccine.

In a preferred embodiment, *in situ* hybridization of labeled lung cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including lung cancer tissue and/or normal tissue, are made. *In situ* hybridization (see, e.g., Ausubel, *supra*) is then performed. When comparing the fingerprints between an individual and a standard, the skilled artisan can make a diagnosis, a prognosis, or a prediction based on the findings. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or may be predictive of outcomes.

In a preferred embodiment, the lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to lung cancer, clinical, pathological, or other information, in terms of long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. Single or multiple

genes may be useful in various combinations. As above, lung cancer probes may be attached to biochips for the detection and quantification of lung cancer sequences in a tissue or patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

5

Assays for therapeutic compounds

In a preferred embodiment, the proteins, nucleic acids, and antibodies as described herein are used in drug screening assays. The lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in drug screening
10 assays or by evaluating the effect of drug candidates on a "gene expression profile" or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Zlokarnik, et al. (1998) Science 279:84-8; Heid (1996) Genome Res. 6:986-94.

15 In a preferred embodiment, the lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified lung cancer proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the lung cancer phenotype or an identified physiological function of a lung cancer protein. As above, this can be done on an individual gene level or
20 by evaluating the effect of drug candidates on a "gene expression profile". In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, *supra*.

Having identified differentially expressed genes herein, a variety of assays may be
25 performed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene with altered regulation in lung cancer, test compounds can be screened for the ability to modulate gene expression or for binding to the lung cancer protein. "Modulation" thus includes an increase or a decrease in gene expression. The preferred amount of modulation will depend on the original change of the
30 gene expression in normal versus tissue undergoing lung cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in lung cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in lung

cancer tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, e.g., through the use of antibodies to the lung cancer protein and standard immunoassays. Proteomics and separation techniques may also allow quantification of expression.

In a preferred embodiment, gene or protein expression monitoring of a number of entities, i.e., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein.

In this embodiment, the lung cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of lung cancer sequences in a particular cell. Alternatively, PCR may be used. Thus, a series, e.g., of microtiter plate, may be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Expression monitoring can be performed to identify compounds that modify the expression of one or more lung cancer-associated sequences, e.g., a polynucleotide sequence set out in the tables. Generally, in a preferred embodiment, a test compound is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate lung cancer, modulate lung cancer proteins, bind to a lung cancer protein, or interfere with the binding of a lung cancer protein and an antibody, substrate, or other binding partner.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes a molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the lung cancer phenotype or the expression of a lung cancer sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments, modulators alter expression profiles of nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses a lung cancer phenotype, e.g., to a normal or non-malignant tissue fingerprint. In another embodiment, a modulator induces a lung cancer phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

In one aspect, a modulator will neutralize the effect of a lung cancer protein. By "neutralize" is meant that activity of a protein and the consequent effect on the cell is inhibited or blocked.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to a lung cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop, et al. (1994) J. Med. Chem. 37(9):1233-1251).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Patent No. 5,010,175, Furka (1991) Pept. Prot. Res. 37:487-493, Houghton, et al. (1991) Nature, 354:84-88), peptoids (PCT Publication No WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs, et al. (1993) Proc. Nat. Acad. Sci. USA 90:6909-6913), vinylogous polypeptides (Hagihara, et al. (1992) J. Amer. Chem. Soc. 114:6568), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann, et

al. (1992) J. Amer. Chem. Soc. 114:9217-9218), analogous organic syntheses of small compound libraries (Chen, et al. (1994) J. Amer. Chem. Soc. 116:2661), oligocarbamates (Cho, et al. (1993) Science 261:1303), and/or peptidyl phosphonates (Campbell, et al. (1994) J. Org. Chem. 59:658). See, generally, Gordon, et al. (1994) J. Med. Chem. 37:1385, nucleic acid libraries (see, e.g., Stratagene, Corp.), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), antibody libraries (see, e.g., Vaughn, et al. (1996) Nature Biotechnology 14(3):309-314, and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang, et al. (1996) Science 274:1520-1522, and U.S. Patent No. 5,593,853), and small organic molecule libraries (see, e.g., benzodiazepines, Baum (1993) C&EN, Jan 18, page 33; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (see, e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. The above devices, with appropriate modification, are suitable for use with the present invention. In addition, numerous combinatorial libraries are themselves commercially available (see, e.g., ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

The assays to identify modulators are amenable to high throughput screening. Preferred assays thus detect modulation of lung cancer gene transcription, polypeptide expression, and polypeptide activity.

High throughput assays for evaluating the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, e.g., U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins,

U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (i.e., in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (see, e.g., Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). These systems typically automate procedures, including sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, e.g., Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

In one embodiment, modulators are proteins, often naturally occurring proteins or fragments of naturally occurring proteins. Thus, e.g., cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of proteins may be made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly useful test compound will be directed to the class of proteins to which the target belongs, e.g., substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that the nucleic acid or peptide consists of essentially random sequences of nucleotides and amino acids, respectively. Since these random peptides (or nucleic acids, discussed below) are often chemically synthesized, they may incorporate a nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. In a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, e.g., of hydrophobic amino acids, hydrophilic residues, 5 sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc.

Modulators of lung cancer can also be nucleic acids, as defined above.

10 As described above generally for proteins, nucleic acid modulating agents may be naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. Digests of procaryotic or eucaryotic genomes may be used as is outlined above for proteins.

In a preferred embodiment, the candidate compounds are organic chemical moieties, a wide variety of which are available in the literature.

15 After a candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence is analyzed. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an *in vitro* transcription 20 with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

In a preferred embodiment, the target sequence is labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as, 25 alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin 30 is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

Nucleic acid assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos.

5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allow formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration, pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein may be accomplished in a variety of ways. Components of the reaction may be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g., albumin, detergents, etc. which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the lung cancer phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, e.g., for diagnostic applications, having identified differentially expressed genes important in a particular state, screens can be performed to identify modulators that alter expression of individual genes. In another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state,

screens are performed to identify agents that bind and/or modulate the biological activity of the gene product, or evaluate genetic polymorphisms.

Genes can be screened for those that are induced in response to a candidate agent. After identifying a modulator based upon its ability to suppress a lung cancer expression pattern leading to a normal expression pattern, or to modulate a single lung cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated lung cancer tissue reveals genes that are not expressed in normal tissue or lung cancer tissue, but are expressed in agent treated tissue. These agent-specific sequences can be identified and used by methods described herein for lung cancer genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the agent induced proteins and used to target novel therapeutics to the treated lung cancer tissue sample.

Thus, in one embodiment, a test compound is administered to a population of lung cancer cells, that have an associated lung cancer expression profile. By "administration" or "contacting" herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (i.e., a peptide) may be put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used.

Once a test compound has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

Thus, e.g., lung cancer or non-malignant tissue may be screened for agents that modulate, e.g., induce or suppress a lung cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on lung cancer activity. By defining such a signature for the lung cancer phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

Measure of lung cancer polypeptide activity, or of lung cancer or the lung cancer phenotype can be performed using a variety of assays. For example, the effects of the test compounds upon the function of the metastatic polypeptides can be measured by examining parameters described above. A suitable physiological change that affects activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of lung cancer associated with tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGMP. In the assays of the invention, mammalian lung cancer polypeptide is typically used, e.g., mouse, preferably human.

Assays to identify compounds with modulating activity can be performed *in vitro*. For example, a lung cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the lung cancer polypeptide levels are determined *in vitro* by measuring the level of protein or mRNA. The level of protein is typically measured using immunoassays such as western blotting, ELISA and the like with an antibody that selectively binds to the lung cancer polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is typically detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using a lung cancer protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or β -gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself can be done. The gene products of differentially expressed

genes are sometimes referred to herein as "lung cancer proteins." The lung cancer protein may be a fragment, or alternatively, be the full length protein to a fragment shown herein.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another
5 embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or isolated
10 gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the lung cancer proteins can be used in the assays.

Thus, in a preferred embodiment, the methods comprise combining a lung cancer
15 protein and a candidate compound, and determining the binding of the compound to the lung cancer protein. Preferred embodiments utilize the human lung cancer protein, although other mammalian proteins may also be used, e.g., for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative lung cancer proteins may be used.

20 Generally, in a preferred embodiment of the methods herein, the lung cancer protein or the candidate agent is non-diffusably bound to an insoluble support, preferably having isolated sample receiving areas (e.g., a microtiter plate, an array, etc.). The insoluble supports may be made of a composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of
25 screening. The surface of such supports may be solid or porous and of a convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents
30 and samples. The particular manner of binding of the composition is typically not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition, and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation

sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

In a preferred embodiment, the lung cancer protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the support and the lung cancer protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

The determination of the binding of the test modulating compound to the lung cancer protein may be done in a number of ways. In a preferred embodiment, the compound is labeled, and binding determined directly, e.g., by attaching all or a portion of the lung cancer protein to a solid support, adding a labeled candidate agent (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as appropriate.

In some embodiments, only one of the components is labeled, e.g., the proteins (or proteinaceous candidate compounds) can be labeled. Alternatively, more than one component can be labeled with different labels, e.g., ^{125}I for the proteins and a fluorophor for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor may be a binding moiety known to bind to the target molecule (i.e., a lung cancer protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at a temperature which facilitates optimal activity, typically between 4 and 40° C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening. Typically

between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by a test compound. Displacement of the competitor is an indication that the test compound is binding to the lung cancer protein and thus is capable of binding to, and potentially modulating, the activity of the lung cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the lung cancer protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is capable of binding to the lung cancer protein.

In a preferred embodiment, the methods comprise differential screening to identify agents that are capable of modulating the activity of the lung cancer proteins. In one embodiment, the methods comprise combining a lung cancer protein and a competitor in a first sample. A second sample comprises a test compound, a lung cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the lung cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the lung cancer protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native lung cancer protein, but cannot bind to modified lung cancer proteins. The structure of the lung cancer protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of a lung cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

Positive controls and negative controls may be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results.

Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g., albumin, detergents, etc. which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in an order that provides for the requisite binding.

In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of a lung cancer protein. The methods comprise adding a test compound, as defined above, to a cell comprising lung cancer proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a lung cancer protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g., hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (e.g., cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

In this way, compounds that modulate lung cancer agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the lung cancer protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

In one embodiment, a method of inhibiting lung cancer cell division is provided. The method comprises administration of a lung cancer inhibitor. In another embodiment, a method of inhibiting lung cancer is provided. The method may comprise administration of a lung cancer inhibitor. In a further embodiment, methods of treating cells or individuals with lung cancer are provided, e.g., comprising administration of a lung cancer inhibitor.

In one embodiment, a lung cancer inhibitor is an antibody as discussed above. In another embodiment, the lung cancer inhibitor is an antisense molecule.

A variety of cell growth, proliferation, viability, and metastasis assays are known to those of skill in the art, as described below.

Soft agar growth or colony formation in suspension

5 Normal cells require a solid substrate to attach and grow. When the cells are transformed, they lose this phenotype and grow detached from the substrate. For example, transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, regenerate normal phenotype and require a solid substrate to attach and
10 grow. Soft agar growth or colony formation in suspension assays can be used to identify modulators of lung cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A therapeutic compound would reduce or eliminate the host cells' ability to grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft.

15 Techniques for soft agar growth or colony formation in suspension assays are described in Freshney (1994) Culture of Animal Cells a Manual of Basic Technique (3rd ed.), herein incorporated by reference. See also, the methods section of Garkavtsev, et al. (1996), *supra*, herein incorporated by reference.

20 *Contact inhibition and density limitation of growth*

Normal cells typically grow in a flat and organized pattern in a petri dish until they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. When cells are transformed, however, the cells are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, the transformed cells grow to a
25 higher saturation density than normal cells. This can be detected morphologically by the formation of a disoriented monolayer of cells or rounded cells in foci within the regular pattern of normal surrounding cells. Alternatively, labeling index with (³H)-thymidine at saturation density can be used to measure density limitation of growth. See Freshney (1994), *supra*. The transformed cells, when transfected with tumor suppressor genes, regenerate a
30 normal phenotype and become contact inhibited and would grow to a lower density.

In this assay, labeling index with (³H)-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected with a lung cancer-associated sequence and are grown for 24 hours at saturation density in

non-limiting medium conditions. The percentage of cells labeling with (^3H)-thymidine is determined autoradiographically. See, Freshney (1994), *supra*.

Growth factor or serum dependence

5 Transformed cells typically have a lower serum dependence than their normal counterparts (see, e.g., Temin (1966) J. Natl. Cancer Inst. 37:167-175; Eagle, et al. (1970) J. Exp. Med. 131:836-879); Freshney, *supra*. This is in part due to release of various growth factors by the transformed cells. Growth factor or serum dependence of transformed host cells can be compared with that of control.

Tumor specific markers levels

Tumor cells release an increased amount of certain factors (hereinafter "tumor specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (see, e.g.,
15 Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" in Mihich (ed. 1985) Biological Responses in Cancer, pp. 178-184). Similarly, Tumor angiogenesis factor (TAF) is released at a higher level in tumor cells than their normal counterparts. See, e.g., Folkman (1992) "Angiogenesis and Cancer" in Sem Cancer Biol.

Various techniques which measure the release of these factors are described in
20 Freshney (1994), *supra*. Also, see, Unkeless, et al. (1974) J. Biol. Chem. 249:4295-4305; Strickland and Beers (1976) J. Biol. Chem. 251:5694-5702; Whur, et al. (1980) Br. J. Cancer 42:305-312; Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" in Mihich (ed. 1985) Biological Responses in Cancer, pp. 178-184; Freshney Anticancer Res. 5:111-130 (1985).

Invasiveness into Matrigel

The degree of invasiveness into Matrigel or some other extracellular matrix constituent can be used as an assay to identify compounds that modulate lung cancer-associated sequences. Tumor cells exhibit a good correlation between malignancy and
30 invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells.

Techniques described in Freshney (1994), *supra*, can be used. Briefly, the level of invasion of host cells can be measured by using filters coated with Matrigel or some other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with ^{125}I and counting the radioactivity on the distal side of the filter or bottom of the dish. See, e.g., Freshney (1984), *supra*.

Tumor growth in vivo

Effects of lung cancer-associated sequences on cell growth can be tested in transgenic or immune-suppressed mice. Knock-out transgenic mice can be made, in which the lung cancer gene is disrupted or in which a lung cancer gene is inserted. Knock-out transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous lung cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous lung cancer gene with a mutated version of the lung cancer gene, or by mutating the endogenous lung cancer gene, e.g., by exposure to carcinogens.

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi, et al. (1989) *Science* 244:1288). Chimeric targeted mice can be derived according to Hogan, et al. (1988) Manipulating the Mouse Embryo: A Laboratory Manual, Cold Spring Harbor Laboratory and Robertson (ed. 1987) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, IRL Press, Washington, D.C.

Alternatively, various immune-suppressed or immune-deficient host animals can be used. For example, genetically athymic "nude" mouse (see, e.g., Giovanella, et al. (1974) *J. Natl. Cancer Inst.* 52:921), a SCID mouse, a thymectomized mouse, or an irradiated mouse (see, e.g., Bradley, et al. (1978) *Br. J. Cancer* 38:263; Selby, et al. (1980) *Br. J. Cancer* 41:52) can be used as a host. Transplantable tumor cells (typically about 10^6 cells) injected into isogenic hosts will produce invasive tumors in a high proportions of cases, while normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing a lung cancer-associated sequences are injected subcutaneously. After a suitable length of time,

preferably 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

5 Polynucleotide modulators of lung cancer

Antisense and RNAi Polynucleotides

In certain embodiments, the activity of a lung cancer-associated protein is downregulated, or entirely inhibited, by the use of antisense or an inhibitory polynucleotide, i.e., a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, e.g., a lung cancer protein mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprehended by this invention so long as they function effectively to hybridize with the lung cancer protein mRNA. See, e.g., Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides. Sense oligonucleotides can, e.g., be employed to block transcription by binding to the antisense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for lung cancer molecules. A preferred antisense molecule is for a lung cancer sequence in the tables, or for a ligand or activator thereof. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein

is described in, e.g., Stein and Cohen (1988) Cancer Res. 48:2659 and van der Krol, et al. (1988) BioTechniques 6:958).

RNA interference is a mechanism to suppress gene expression in a sequence specific manner. See, e.g., Brumelkamp, et al. (2002) Scienceexpress (21 March 2002); Sharp (1999) Genes Dev. 13:139-141; and Cathew (2001) Curr. Op. Cell Biol. 13:244-248. In mammalian cells, short, e.g., 21 nt, double stranded small interfering RNAs (siRNA) have been shown to be effective at inducing an RNAi response. See, e.g., Elbashir, et al. (2001) Nature 411:494-498. The mechanism may be used to downregulate expression levels of identified genes, e.g., treatment of or validation of relevance to disease.

10

Ribozymes

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of lung cancer-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (see, e.g., Castanotto, et al. (1994) Adv. in Pharmacology 25: 289-317 for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel, et al. (1990) Nucl. Acids Res. 18:299-304; European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art (see, e.g., WO 94/26877; Ojwang, et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Yamada, et al. (1994) Human Gene Therapy 1:39-45; Leavitt, et al. (1995) Proc. Natl. Acad. Sci. USA 92:699-703; Leavitt, et al. (19994) Human Gene Therapy 5:1151-120; and Yamada, et al. (1994) Virology 205: 121-126).

Polynucleotide modulators of lung cancer may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of lung cancer may be introduced into a cell containing the target nucleic acid sequence, e.g., by

formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

Thus, in one embodiment, methods of modulating lung cancer in cells or organisms are provided. In one embodiment, the methods comprise administering to a cell an anti-lung cancer antibody that reduces or eliminates the biological activity of an endogenous lung cancer protein. Alternatively, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding a lung cancer protein. This may be accomplished in any number of ways. In a preferred embodiment, e.g., when the lung cancer sequence is down-regulated in lung cancer, such state may be reversed by increasing the amount of lung cancer gene product in the cell. This can be accomplished, e.g., by overexpressing the endogenous lung cancer gene or administering a gene encoding the lung cancer sequence, using known gene-therapy techniques. In a preferred embodiment, the gene therapy techniques include the incorporation of the exogenous gene using enhanced homologous recombination (EHR), e.g., as described in PCT/US93/03868, hereby incorporated by reference in its entirety. Alternatively, e.g., when the lung cancer sequence is up-regulated in lung cancer, the activity of the endogenous lung cancer gene is decreased, e.g., by the administration of a lung cancer antisense or RNAi nucleic acid.

In one embodiment, the lung cancer proteins of the present invention may be used to generate polyclonal and monoclonal antibodies to lung cancer proteins. Similarly, the lung cancer proteins can be coupled, using standard technology, to affinity chromatography columns. These columns may then be used to purify lung cancer antibodies useful for production, diagnostic, or therapeutic purposes. In a preferred embodiment, the antibodies are generated to epitopes unique to a lung cancer protein; that is, the antibodies show little or no cross-reactivity to other proteins. The lung cancer antibodies may be coupled to standard affinity chromatography columns and used to purify lung cancer proteins. The antibodies may also be used as blocking polypeptides, as outlined above, since they will specifically bind to the lung cancer protein.

30 **Methods of identifying variant lung cancer-associated sequences**

Without being bound by theory, expression of various lung cancer sequences is correlated with lung cancer. Accordingly, disorders based on mutant or variant lung cancer genes may be determined. In one embodiment, the invention provides methods for

identifying cells containing variant lung cancer genes, e.g., determining all or part of the sequence of at least one endogenous lung cancer genes in a cell. In a preferred embodiment, the invention provides methods of identifying the lung cancer genotype of an individual, e.g., determining all or part of the sequence of at least one lung cancer gene of the individual.

- 5 This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced lung cancer gene to a known lung cancer gene, i.e., a wild-type gene.

The sequence of all or part of the lung cancer gene can then be compared to the
10 sequence of a known lung cancer gene to determine if any differences exist. This can be done using known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the lung cancer gene of the patient and the known lung cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

- 15 In a preferred embodiment, the lung cancer genes are used as probes to determine the number of copies of the lung cancer gene in the genome.

In another preferred embodiment, the lung cancer genes are used as probes to determine the chromosomal localization of the lung cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when
20 chromosomal abnormalities such as translocations, and the like are identified in the lung cancer gene locus.

Administration of pharmaceutical and vaccine compositions

In one embodiment, a therapeutically effective dose of a lung cancer protein or
25 modulator thereof, is administered to a patient. By "therapeutically effective dose" herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (e.g., Ansel, et al. (1992) Pharmaceutical Dosage Forms and Drug Delivery; Lieberman, Pharmaceutical Dosage Forms (vols. 1-3), Dekker, ISBN 0824770846,
30 082476918X, 0824712692, 0824716981; Lloyd (1999) The Art, Science and Technology of Pharmaceutical Compounding; and Pickar (1999) Dosage Calculations). Adjustments for lung cancer degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration,

drug interaction and the severity of the condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art.

A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals. Thus the methods are applicable to both human therapy and veterinary applications. In the preferred embodiment the patient is a mammal, preferably a primate, and in the most preferred embodiment the patient is human.

The administration of the lung cancer proteins and modulators thereof of the present invention can be done in a variety of ways, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly, intrapulmonary, vaginally, rectally, or intraocularly. In some instances, e.g., in the treatment of wounds and inflammation, the lung cancer proteins and modulators may be directly applied as a solution or spray.

The pharmaceutical compositions of the present invention comprise a lung cancer protein in a form suitable for administration to a patient. In the preferred embodiment, the pharmaceutical compositions are in a water soluble form, such as being present as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. "Pharmaceutically acceptable acid addition salt" refers to those salts that retain the biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose,

lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol.

The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that lung cancer protein modulators (e.g., antibodies, antisense constructs, ribozymes, small organic molecules, etc.) when administered orally, should be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by packaging the molecule(s) in an appropriately resistant carrier, such as a liposome or a protection barrier. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise a lung cancer protein modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, e.g., sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs (e.g., Remington's Pharmaceutical Science (15th ed., 1980) and Hardman, et al. (eds. 1996) Goodman and Gilman: The Pharmacological Basis of Therapeutics).

Thus, a typical pharmaceutical composition for intravenous administration would be about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally administrable compositions will be known or apparent to those skilled in the art, e.g., Remington's Pharmaceutical Science and Goodman and Gilman, The Pharmacological Basis of Therapeutics, *supra*.

The compositions containing modulators of lung cancer proteins can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g., a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient. An amount of modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a "prophylactically effective dose." The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, administration route, efficiency, etc. Such prophylactic treatments may be used, e.g., in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer based, at least in part, upon gene expression profiles. Vaccine strategies may be used, in either a DNA vaccine form, or protein vaccine.

It will be appreciated that the present lung cancer protein-modulating compounds can be administered alone or in combination with additional lung cancer modulating compounds or with other therapeutic agent, e.g., other anti-cancer agents or treatments.

In numerous embodiments, one or more nucleic acids, e.g., polynucleotides comprising nucleic acid sequences set forth in the tables, such as antisense or RNAi polynucleotides or ribozymes, will be introduced into cells, *in vitro* or *in vivo*. The present invention provides methods, reagents, vectors, and cells useful for expression of lung cancer-associated polypeptides and nucleic acids using *in vitro* (cell-free), *ex vivo*, or *in vivo* (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Many procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors and other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (see, e.g.,

Berger and Kimmel; Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 (Berger), Ausubel, et al. (eds. 1999) Current Protocols (supplemented through 1999), and Sambrook, et al. (1989) Molecular Cloning - A Laboratory Manual (2nd ed., Vol. 1-3).

- 5 In a preferred embodiment, lung cancer proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly, lung cancer genes (including both the full-length sequence, partial sequences, or regulatory sequences of the lung cancer coding regions) can be administered in a gene therapy application. These lung cancer genes can include antisense or inhibitory applications, e.g., as inhibitory RNA or gene
10 therapy (e.g., for incorporation into the genome) or as antisense compositions.

- Lung cancer polypeptides and polynucleotides can also be administered as vaccine compositions to stimulate HTL, CTL, and antibody responses.. Such vaccine compositions can include, e.g., lipidated peptides (see, e.g., Vitiello, et al. (1995) J. Clin. Invest. 95:341), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres
15 (see, e.g., Eldridge, et al. (1991) Molec. Immunol. 28:287-294; Alonso, et al. (1994) Vaccine 12:299-306; Jones, et al. (1995) Vaccine 13:675-681), peptide compositions contained in immune stimulating complexes (ISCOMS) (see, e.g., Takahashi, et al. (1990) Nature 344:873-875; Hu, et al. (1998) Clin Exp Immunol. 113:235-243), multiple antigen peptide systems (MAPs) (see, e.g., Tam (1988) Proc. Natl. Acad. Sci. U.S.A. 85:5409-5413; Tam
20 (1996) J. Immunol. Methods 196:17-32), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery vectors (Perkus, et al., p. 379 In: Kaufmann (ed. 1996) Concepts in vaccine development; Chakrabarti, et al. (1986) Nature 320:535; Hu, et al. (1986) Nature 320:537; Kieny, et al. (1986) AIDS Bio/Technology 4:790; Top, et al. (1971) J. Infect. Dis. 124:148; Chanda, et al.
25 (1990) Virology 175:535), particles of viral or synthetic origin (see, e.g., Kofler, et al. (1996) J. Immunol. Methods 192:25; Eldridge, et al. (1993) Sem. Hematol. 30:16; Falo, et al. (1995) Nature Med. 7:649), adjuvants (Warren, et al. (1986) Annu. Rev. Immunol. 4:369; Gupta, et al. (1993) Vaccine 11:293), liposomes (Reddy, et al. (1992) J. Immunol. 148:1585; Rock (1996) Immunol. Today 17:131), or, naked or particle absorbed cDNA (Ulmer, et al. (1993)
30 Science 259:1745; Robinson, et al. (1993) Vaccine 11:957; Shiver, et al., p. 423 In: Kaufmann (ed. 1996) Concepts in vaccine development; Cease and Berzofsky (1994) Annu. Rev. Immunol. 12:923 and Eldridge, et al. (1993) Sem. Hematol. 30:16). Toxin-targeted

delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Vaccine compositions often include adjuvants. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, e.g., Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Vaccines can be administered as nucleic acid compositions wherein DNA or RNA encoding one or more of the polypeptides, or a fragment thereof, is administered to a patient. This approach is described, for instance, in Wolff, et. al. (1990) Science 247:1465 as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivacaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (see, e.g., U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus, e.g., as a vector to express nucleotide sequences that encode lung cancer polypeptides or polypeptide fragments. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover, et al. (1991) Nature 351:456-460. A wide variety of other vectors useful for therapeutic administration or immunization e.g., adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the

like, will be apparent to those skilled in the art from the description herein (see, e.g., Shata, et al. (2000) Mol Med Today 6:66-71; Shedlock, et al. (2000) J. Leukoc. Biol. 68:793-806; Hipp, et al. (2000) In Vivo 14:571-85).

Methods for the use of genes as DNA vaccines are well known, and include placing a lung cancer gene or portion of a lung cancer gene under the control of a regulatable promoter or a tissue-specific promoter for expression in a lung cancer patient. The lung cancer gene used for DNA vaccines can encode full-length lung cancer proteins, but more preferably encodes portions of the lung cancer proteins including peptides derived from the lung cancer protein. In one embodiment, a patient is immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from a lung cancer gene. For example, lung cancer-associated genes or sequence encoding subfragments of a lung cancer protein are introduced into expression vectors and tested for their immunogenicity in the context of Class I MHC and an ability to generate cytotoxic T cell responses. This procedure provides for production of cytotoxic T cell responses against cells which present antigen, including intracellular epitopes.

In a preferred embodiment, DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the lung cancer polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are available.

In another preferred embodiment lung cancer genes find use in generating animal models of lung cancer. When the lung cancer gene identified is repressed or diminished in metastatic tissue, gene therapy technology, e.g., wherein antisense or inhibitory RNA directed to the lung cancer gene will also diminish or repress expression of the gene. Animal models of lung cancer find use in screening for modulators of a lung cancer-associated sequence or modulators of lung cancer. Similarly, transgenic animal technology including gene knockout technology, e.g., as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence or increased expression of the lung cancer protein. When desired, tissue-specific expression or knockout of the lung cancer protein may be necessary.

It is also possible that the lung cancer protein is overexpressed in lung cancer. As such, transgenic animals can be generated that overexpress the lung cancer protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene.

Animals generated by such methods will find use as animal models of lung cancer and are additionally useful in screening for modulators to treat lung cancer.

Kits for Use in Diagnostic and/or Prognostic Applications

5 For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In diagnostic and research applications such kits may include at least one of the following: assay reagents, buffers, lung cancer-specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, RNAi, dominant negative lung cancer polypeptides or polynucleotides, small molecule inhibitors of
10 lung cancer-associated sequences, etc. A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

In addition, the kits may include instructional materials containing instructions (e.g., protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. A medium
15 capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

The present invention also provides for kits for screening for modulators of lung
20 cancer-associated sequences. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise one or more of the following materials: a lung cancer-associated polypeptide or polynucleotide, reaction tubes, and instructions for testing lung cancer-associated activity. Optionally, the kit contains biologically active lung cancer protein. A wide variety of kits and components can be prepared according to the present
25 invention, depending upon the intended user of the kit and the particular needs of the user. Diagnosis would typically involve evaluation of a plurality of genes or products. The genes typically will be selected based on correlations with important parameters in disease which may be identified in historical or outcome data.

EXAMPLES

Example 1: Gene Chip Analysis

- 5 Molecular profiles of various normal and cancerous tissues were determined and analyzed using gene chips. RNA was isolated and gene chip analysis was performed as described (Glynnne, et al. (2000) Nature 403:672-676; Zhao, et al. (2000) Genes Dev. 14:981-993).

Tables 1A and 1B were previously filed on April 18, 2001 in USSN 60/284,770 (18501-001500US) and on November 29, 2001 in USSN 60/334,370 (18501-001520US)

5	Table 1A				70% chron/90% NL	70% SQUAD/90% NL
	Pkey	ExAccn	UnigeneID	Unigene Title		
	100134	D13264	Hs.49	macrophage scavenger receptor 1	1.61	0.74
	100780	HG3731-HT4001		Immunoglobulin Heavy Chain, V α 1c Reg	2.68	3.28
10	100971	J02874	Hs.83213	fatty acid binding protein 4; adipocyte	1.96	0.14
	101088	L05568	Hs.553	solute carrier family 6 (neurotransmitter)	0.79	0.07
	101102	L07594	Hs.79059	transforming growth factor; beta receptor	2.55	1
	101168	L15398	Hs.211559	G protein-coupled receptor kinase 5	0.88	0.27
	101277	L38486	Hs.118223	microfibrillar-associated protein 4	0.89	0.26
	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do	0.59	0.29
15	101336	L49169	Hs.75678	FBJ murine osteosarcoma viral oncogene h	1.15	0.41
	101345	L76380	Hs.152175	calctonin receptor-like	0.81	0.31
	101678	M62505	Hs.2161	complement component 5 receptor 1 (C5a I	1.31	0.77
	101764	M80563	Hs.81256	S100 calcium-binding protein A4 (calcium	1.44	0.82
	101771	M81750	Hs.153837	myeloid cell nuclear differentiation ant	0.96	0.45
20	101842	M93221	Hs.75182	mannose receptor; C type 1	1.27	0.37
	102283	U31384	Hs.83381	guanine nucleotide binding protein 11	1.04	0.3
	102363	U39447	Hs.198241	amine oxidase; copper containing 3 (vasc	0.96	0.26
	102507	U52154	Hs.193044	potassium inwardly-rectifying channel; s	2.81	3.45
	102698	U75272	Hs.1867	progastricsin (pepsinogen C)	0.95	0.23
25	103025	X54131	Hs.123641	protein tyrosine phosphatase; receptor t	1.62	0.21
	103280	X79881	Hs.76205	cadherin 5; VE-cadherin (vascular epithe	0.9	0.41
	103496	Y09267	Hs.132821	flavin containing monooxygenase 2	1.27	0.49
	103541	Z11697	Hs.79197	CD83 antigen (activated B lymphocytes; I	1.86	1
	103554	Z18951	Hs.74034	caveolin 1; caveolae protein; 22kD	1.27	0.47
30	104212	AB002298	Hs.173035	KIAA0300 protein	1.17	0.16
	104691	AA011176	Hs.37744	ESTs	1.08	0.35
	104825	AA035613	Hs.141883	ESTs	0.75	0.27
	104857	AA043219	Hs.19058	ESTs	2.6	3.3
	104865	AA045136	Hs.22575	ESTs	1.23	0.49
35	104989	AA102099	Hs.118615	ESTs	0.63	0.32
	105729	AA292694	Hs.3807	ESTs; Weakly similar to PHOSPHOLEMMAN PR	0.86	0.34
	105847	AA398606	Hs.32241	ESTs	1.32	0.4
	105894	AA400979	Hs.25691	calctonin receptor-like receptor activi	0.78	0.28
40	106490	AA451861	Hs.115537	ESTs; Weakly similar to dipeptidase prec	1.2	0.47
	106536	AA453997	Hs.23804	ESTs	0.82	0.15
	106505	AA457718	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr	0.99	0.07
	106667	AA461086	Hs.16578	ESTs	1.17	0.4
	106773	AA478109	Hs.188833	ESTs	1.46	0.43
	106797	AA478962	Hs.169943	ESTs	1.18	0.32
45	106844	AA485055	Hs.158213	sperm associated antigen 6	0.98	0.51
	106870	AA487576	Hs.26530	serum deprivation response (phosphatidyl	1.05	0.14
	106954	AA496980	Hs.204038	ESTs	1.25	0.33
	107054	AA600150	Hs.14356	ESTs	1.11	0.4
	107292	T30407	Hs.4789	ESTs; Weakly similar to oxidative-stress	1.07	2.58
50	107994	AA036811	Hs.165030	ESTs	0.7	0.21
	107997	AA037388	Hs.82223	Human DNA sequence from clone 141H5 on c	1.02	0.48
	108041	AA041552	Hs.61957	ESTs	1.44	0.51
	108087	AA045709	Hs.40545	ESTs	1.98	1
	108382	AA074885	Hs.67726	macrophage receptor with collagenous str	1.52	0.72
55	108435	AA078787	Hs.194101	ESTs	2.53	1.53
	108480	AA081093	Hs.68055	ESTs	1.56	0.48
	109252	AA194830	Hs.85944	ESTs	2.69	3.18
	109550	F01534	Hs.26981	ESTs	1.19	0.65
60	109613	F03031	Hs.27519	ESTs	1.01	0.29
	109837	H00656	Hs.29792	ESTs	0.81	0.15
	109893	H04768	Hs.30494	ESTs	1.44	0.32
	109984	H09594	Hs.10299	ESTs	0.62	0.14
	110099	H16568	Hs.23748	ESTs	1.01	0.28
	110837	N30796	Hs.17424	ESTs; Weakly similar to semaphorin F [H.	1.1	0.22
65	111247	N69825	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	1.26	0.26
	111341	N80935	Hs.22483	ESTs	1.57	0.52
	111510	R07856	Hs.16355	ESTs	3.96	1
	111737	R25410	Hs.9218	ESTs	0.97	0.24
	113195	T57112		***yc20g11.s1 Stratogene lung (#937210)	1.22	0.35
70	113238	T62979	Hs.189813	ESTs	2.27	0.45
	113540	T90496	Hs.16757	ESTs	1.06	0.22
	113552	T90889	Hs.16026	ESTs	1.16	0.42
	113606	T93093	Hs.17125	ESTs	1.48	0.7
	113695	T95965	Hs.17948	ESTs	1.54	0.28
75	113946	W84753	Hs.37896	ESTs	1.79	0.72
	114251	Z39898	Hs.21948	ESTs	1.95	0.25
	114359	Z41589	Hs.153483	ESTs; Moderately similar to H1 chloride	1.42	0.13
	115230	AA278300	Hs.182980	ESTs	2.62	0.42
	115279	AA279760	Hs.63671	ESTs	1.79	0.91
80	115566	AA398083	Hs.43977	ESTs	0.86	0.2
	115965	AA446661	Hs.173233	ESTs	0.79	0.04
	116166	AA461556	Hs.202949	KIAA1102 protein	2.29	0.68
	116279	AA486073	Hs.57362	ESTs	2.27	0.78
	117023	H88157	Hs.41105	ESTs	1.36	0.16

	117209	H99959	Hs.42768	ESTs	1.46	0.48
	118901	N90719	Hs.94445	ESTs	1.51	1
	118981	N93839	Hs.39288	ESTs	1.34	0.48
	119073	R32894	Hs.45514	v-ets avian erythroblastosis virus E26 o	1.14	0.27
5	119221	R98105		***yr30g11.s1 Soares fetal liver splean	1.32	0.53
	119824	W74536	Hs.184	advanced glycosylation end product-speci	1	0.19
	119861	W80715		ESTs; Moderately similar to IIII ALU SUB	1.83	0.45
	120041	W92775	Hs.59368	ESTs	1.23	0.55
10	120132	Z38839	Hs.125019	ESTs; Highly similar to KIAA0886 protein	0.91	0.37
	120467	AA251579	Hs.187628	ESTs	1.87	1.91
	121314	AA402799	Hs.182538	ESTs	1.3	0.31
	121643	AA417078	Hs.193767	ESTs	2.31	0.68
	121690	AA418074	Hs.110286	ESTs	1.47	0.51
	122633	AA454080	Hs.34853	inhibitor of DNA binding 4; dominant neg	1.31	0.63
15	123978	C20653	Hs.170278	ESTs	1.52	0.32
	124214	H58608	Hs.151323	ESTs	0.93	0.35
	124357	N22401		***yw37g07.s1 Morton Fetal Cochlea Homo	1.29	1
	124438	N40188	Hs.102550	ESTs	1.36	0.7
20	125167	W45550	Hs.102541	ESTs	1.45	0.69
	125174	W51835	Hs.231082	EST	3.07	3.76
	125422	AA903229	Hs.153717	ESTs	1.34	0.3
	125561	AA17667	Hs.22978	ESTs	1.89	0.63
	125831	D60988		***HUM145809B Clontech human fetal brain	0.94	0.36
25	127002	R35380	Hs.24979	ESTs	3.02	4.06
	127307	AA369367	Hs.126712	ESTs; Weakly similar to pL2 hypotheticala	1.01	0.69
	127609	AA622559	Hs.150318	ESTs	1.21	0.32
	127959	AI302471	Hs.124292	ESTs	2.5	1
	128458	D52193	Hs.56340	ESTs	1.13	0.33
30	128524	AA479209	Hs.102647	ESTs	1.45	0.58
	128789	AA486567	Hs.105695	ESTs	1.1	0.34
	128798	AF014958	Hs.105938	chemokine (C-C motif) receptor-like 2	1.16	0.55
	128952	R51076	Hs.107361	ESTs; Highly similar to Rap2 interacting	2.04	2.4
	129057	X62466	Hs.214742	CDW52 antigen (CAMPATH-1 antigen)	1.77	0.73
35	129210	AA401654	Hs.202949	KIAA1102 protein	1.11	0.36
	129240	W24360	Hs.237868	interleukin 7 receptor	0.91	0.41
	129402	T63781		***yc21g01.s1 Stratagene lung (#937210)	1.36	0.43
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1	0.67	0.08
	129593	AA487015	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	1.3	0.42
40	129526	AA447410	Hs.11712	ESTs; Weakly similar to IIII ALU SUBFAM1	1.28	0.46
	129599	AA458578	Hs.12017	KIAA0439 protein; homolog of yeast ublqu	1.58	1
	129898	N48595	Hs.13256	ESTs	1.13	0.53
	129958	L20591	Hs.1378	annexin A3	0.81	0.31
45	130273	U59914	Hs.153863	MAD (mothers against decapentaplegic; Dr	0.59	0.22
	130655	N92934	Hs.17409	cysteine-rich protein 1 (intestinal)	1.44	0.76
	130657	T94452	Hs.201591	ESTs	0.96	0.42
	131061	N64328	Hs.22567	ESTs; Moderately similar to HYPOTHETICAL	1.51	0.45
	131066	F09006	Hs.22588	ESTs	0.97	0.37
	131263	R38334	Hs.24950	regulator of G-protein signalling 5	2.34	2.82
50	131589	U52100	Hs.29191	epithelial membrane protein 2	1.2	0.62
	131686	AA157428	Hs.30687	Grb2-associated binder 2	0.95	0.38
	131751	H18335	Hs.31562	ESTs	1.47	0.52
	132430	T23630	Hs.258675	EST	1.86	2.09
	132476	N67192	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regi	1.73	0.58
55	132836	F09557	Hs.57929	slit (Drosophila) homolog 3	0.91	0.29
	133120	X64559	Hs.65424	tetranectin (plasminogen-binding protein	0.82	0.2
	133468	D45370	Hs.74120	adipose specific 2	1.29	0.48
	133565	H57056	Hs.204831	ESTs	2.25	0.57
	133651	U97105	Hs.173381	dihydropyrimidinase-like 2	1.65	0.62
60	133835	AA059489	Hs.76540	ESTs; Highly similar to RGC-32 [R.norveg	1.16	0.34
	133978	W73859	Hs.78061	transcription factor 21	0.79	0.27
	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec	0.99	0.28
	134299	AA487558	Hs.8135	ESTs	1.02	0.46
	134300	U81984	Hs.166082	endothelial PAS domain protein 1	0.88	0.42
65	134323	AA028976	Hs.8175	Homo sapiens mRNA; cDNA DKFZp564M0763 (f	1.19	0.27
	134343	D50683	Hs.82028	transforming growth factor; beta recepto	1.21	0.67
	134417	D87969	Hs.82921	solute carrier family 35 (CMP-sialic ac	1.28	1
	134561	U76421	Hs.85302	adenosine deaminase; RNA-specific; B1 (h	2.12	0.55
	134624	W67147	Hs.8700	deleted in liver cancer 1	2.35	2.74
70	134696	H88354	Hs.8861	ESTs	1.35	0.33
	134749	L10955	Hs.89485	carbonic anhydrase IV	0.89	0.2
	134786	L06139	Hs.89640	TEK tyrosine kinase; endothelial (venous	0.48	0.21
	134869	T35288	Hs.90421	ESTs; Moderately similar to IIII ALU SUB	2.14	2.64
	135346	M21056	Hs.992	phospholipase A2; group IB (pancreas)	0.63	0.13
75	100113	D00591	Hs.84746	Chromosome condensation 1	1	2.15
	100147	D13665	Hs.136348	Homo sapiens mRNA for osteoblast specifi	0.5	2
	100280	D42085	Hs.155314	KIAA0095 gene product	1.02	1.39
	100335	D63391	Hs.6793	platelet-activating factor acetylhydrola	1	5.58
	100360	D78335	Hs.75939	Uridine monophosphate kinase	0.91	2.04
	100372	D79997	Hs.184339	KIAA0175 gene product	0.75	2.03
80	100486	HG1112-HT1112		TIGR: ras-like protein TC4	1.09	1.93
	100559	HG2197-HT2267		*collagen, type VII, alpha 1*	0.97	3.6
	100576	HG2290-HT2386		*calitonin/alpha-CGRP, alt. transcript	1	1
	100668	HG2981-HT3938		*TIGR: CD44 (epican, alt. transcript 12	0.85	1.9
	100906	HG4716-HT5158		Guanosine 5'-Monophosphate Synthase	1.18	2.29
85	100930	HG721-HT4827		*TIGR: placental protein 14, endometrial	1	1.45

	100960	J00124	Hs.117729	keratin 14 (epidermolysis bullosa simple	0.84	2.6
	101031	J05070	Hs.151738	*Matrix metalloproteinase 9 (gelatinase	0.77	1.52
	101111	L08424	Hs.1619	Achaete-scute complex (Drosophila) homol	1	1
	101124	L10343	Hs.112341	*Protease inhibitor 3, skin-derived (SKA	0.62	2.67
5	101175	L18920	Hs.36980	*Melanoma antigen, family A, 2	1	1
	101204	L24203	Hs.82237	Ataxia-telangiectasia group D-associated	0.74	4.1
	101431	M19888	Hs.1076	Small proline-rich protein 18 (cornifin)	0.85	2.51
	101448	M21389	Hs.195850	keratin 5 (epidermolysis bullosa simplex	0.61	8.83
	101511	M27826	Hs.267319	Endogenous retroviral protease	1.03	1.13
10	101526	M29540	Hs.220529	Carcinoembryonic antigen-related cell ad	1.07	4.61
	101548	M31328	Hs.71642	*Guanine nucleotide binding protein (G p	0.97	1.13
	101625	M57293		*Human parathyroid hormone-related pepti	1	1
	101649	M60047	Hs.1690	Heparin-binding growth factor binding pr	1	2.7
	101724	M69225	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	1	8.98
15	101748	M76482	Hs.1925	Desmoglein 3 (pemphigus vulgaris antigen	1	2.78
	101759	M80244	Hs.184601	*Solute carrier family 7 (cationic amino	1.07	2.45
	101804	M86699	Hs.169840	TTK protein kinase	1	1
	101806	M86757	Hs.112408	S100 calcium-binding protein A7 (psorias	0.74	1.76
20	101809	M86849		*Homo sapiens connexin 26 (GJB2) mRNA, c	1	7
	101845	M93426	Hs.78867	*Protein tyrosine phosphatase, receptor-	1	1
	101851	M94250	Hs.82045	Midkine (neurite growth-promoting factor	1.13	2.6
	102083	U10323	Hs.75117	*Interleukin enhancer binding factor 2,	1.03	1.61
	102154	U17760	Hs.75517	*Laminin, beta 3 (nlcan (125kD), kalini	0.94	3.62
	102193	U20758	Hs.313	secreted phosphoprotein 1 (osteopontin;	0.34	4.59
25	102305	U33286	Hs.90073	chromosome segregation 1 (yeast homolog)	1.45	2.97
	102348	U37519	Hs.87539	Aldehyde dehydrogenase 8	0.52	2.25
	102581	U61145	Hs.77256	Enhancer of zeste (Drosophila) homolog 2	0.91	2.46
	102610	U65011	Hs.30743	Preferentially expressed antigen in mela	1	3.88
30	102623	U66083	Hs.37110	*Melanoma antigen, family A, 9 (MAGE-9)	1	1
	102669	U71207	Hs.29279	Eyes absent (Drosophila) homolog 2	1	1
	102696	U74612	Hs.239	Forkhead box M1	1.06	2.77
	102829	U91618	Hs.80962	Neurotensin	1	1
	102888	X04741	Hs.76118	Ubiquitin carboxyl-terminal esterase L1	1.13	2.59
35	102913	X07696	Hs.80342	keratin 15	0.7	4.72
	102915	X07820	Hs.2258	Matrix Metalloproteinase 10 (Stromelysin	1.15	3.35
	102963	X15943	Hs.37058	*Calcitonin/calcitonin-related polypepti	1	1
	103021	X53587	Hs.85286	*Integrin, beta 4	1.38	2.34
	103036	X54925	Hs.83169	Matrix metalloproteinase 1 (interstitial c	1	14.93
40	103058	X57348	Hs.184510	Stratifin	1.25	4.17
	103060	X57766	Hs.155324	matrix metalloproteinase 11 (stromelysin	1	1.72
	103119	X63629	Hs.2877	*Cadherin 3, P-cadherin (placental)	1.16	7.38
	103206	X72755	Hs.77357	monokine induced by gamma interferon	0.71	1.48
	103242	X76342	Hs.389	*Alcohol dehydrogenase 7 (class IV), mu	1	1
45	103312	X82693	Hs.3185	*Lymphocyte antigen 6 complex, locus D;	0.92	1.28
	103478	Y07755	Hs.38991	S100 calcium-binding protein A2	1.05	5.81
	103558	Z19574	Hs.2785	keratin 17	0.65	6.68
	103576	Z26317	Hs.2631	Desmoglein 2	0.79	1.73
	103587	Z29083	Hs.82128	ST4 Oncofetal antigen	1	3.93
50	103594	Z31560	Hs.816	*STRY (sex determining region Y)-box 2, p	0.71	7.23
	103768	AA089997		*ESTs, Highly similar to Integral membra	0.99	1.8
	104158	AA454908	Hs.8127	KIAA0144 gene product	0.98	1.29
	104558	R56678	Hs.88959	Human DNA sequence from clone 957N21 on	1.23	7.23
	104589	AA010665		ESTs	0.96	2.11
55	104733	AA019498	Hs.23071	ESTs	1.18	1.88
	104906	AA055809	Hs.26802	Protein kinase domains containing protel	1.11	3.15
	104978	AA088458	Hs.19322	ESTs; Weakly similar to !!! ALU SUBFAM1	1.64	2.89
	105012	AA116036	Hs.9329	*Homo sapiens mRNA for fls353, complete	1.19	3.91
	105175	AA186804	Hs.25740	ESTs; Weakly similar to unknown [S.cerev	0.9	4.63
60	105263	AA227926	Hs.6682	ESTs	0.95	2.87
	105298	AA233459	Hs.26369	ESTs	1	1.13
	105312	AA233854	Hs.23348	S-phase kinase-associated protein 2 (p45	1.32	3.01
	105719	AA291644	Hs.36793	Hypothetical protein FLJ23188	1.28	2.31
	105743	AA293300	Hs.9598	ESTs	1	1
65	106012	AA411621	Hs.8895	ESTs; same as BFH67	0.94	2.04
	106231	AA429571	Hs.38002	KIAA1355 protein	1.04	1.5
	106540	AA454607	Hs.38114	Hypothetical protein FLJ11100	1.26	2.26
	106575	AA456039	Hs.105421	ESTs	1	2
70	106632	AA459897	Hs.11950	GPI-anchored metastasis-associated prote	0.87	1.32
	106727	AA465342	Hs.34045	Hypothetical protein FLJ20764	0.87	1.59
	106906	AA490237	Hs.222024	Transcription factor BMAL2 (cycle-like f	0.61	1.6
	107059	AA608545	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re	0.48	2.67
	107104	AA609786	Hs.15243	Nucleolar protein 1 (120kD)	1.01	1.44
	107151	AA621169	Hs.8687	ESTs; procollagen I-N proteinase	0.97	2.89
	107284	S74039	Hs.291904	Accessory proteins BAP31/BAP29	1.15	3.65
75	107901	AA026418	Hs.91539	ESTs	0.72	3.44
	107922	AA028028	Hs.61460	Ig superfamily receptor LNIR precursor	1	2.48
	107932	AA029317	Hs.18878	Hypothetical protein FLJ21620	1	1
	108695	AA121315	Hs.70823	KIAA1077 protein	0.91	3.53
80	108857	AA133250	Hs.62180	ESTs	1	1
	108860	AA133334	Hs.129911	ESTs	0.73	7.3
	108990	AA152296	Hs.72045	ESTs	1	1
	109166	AA179845	Hs.73625	*RAB6 interacting, kinesin-like (rabkine	1	4.55
	109424	AA227919	Hs.85962	Hyaluronan synthase 3	1	1.28
	109565	F05012	Hs.27027	Hypothetical protein DKFZp762H1311	1.42	2
85	109970	H09281	Hs.13234	ESTs	1.13	2.16

	110015	H10998	Hs.7164	A disintegrin and metalloproteinase domain	0.84	1.95
	110155	H18957	Hs.4213	ESTs	0.94	1.41
	110561	H59617	Hs.5199	HSPC150 protein similar to ubiquitin-con	0.91	3.18
5	111223	N68921	Hs.34806	ESTs; Weakly similar to neogenin [H.sapi	0.91	3.13
	111345	N89820	Hs.14559	Hypothetical protein FLJ10540	1	1.25
	111876	R38239	Hs.293246	*ESTs, Weakly similar to putative p150 [0.83	1.27
	111902	R39191	Hs.109445	KIAA1020 protein	0.91	0.91
	112244	R51309	Hs.70823	KIAA1077 protein	0.77	3.01
10	112973	T17271		*cDNA FLJ13308 fis, clone OVARC1001435,	1	1
	112989	T23482	Hs.89981	"Diacylglycerol kinase, zeta (104kD)"	0.55	1.03
	113047	T25867	Hs.7549	ESTs	0.87	2
	113095	T40920	Hs.126733	ESTs	1	1
	113531	T90345	Hs.16740	Hypothetical protein FLJ11035	0.42	1.44
15	113970	W86748	Hs.8109	ESTs	1.17	1.73
	114346	Z41450	Hs.130489	*ATPase, aminophospholipid transporter-I	0.86	0.82
	114407	AA010188	Hs.103305	ESTs	0.8	1.88
	114471	AA028074	Hs.104613	RP42 homolog	1.06	1.34
	114509	AA043551	Hs.101799	KIAA1350 protein	1.82	2.32
	115060	AA253214	Hs.198249	*Gap junction protein, beta 5 (connexin	0.79	1.49
20	115091	AA255900	Hs.184523	KIAA0965 protein	0.72	1.92
	115123	AA256642	Hs.236894	*ESTs, High sim to LRP1_hu low density I	0.59	1.97
	115291	AA279943	Hs.122579	ESTs	1	1.25
	115506	AA292537	Hs.45207	Hypothetical protein KIAA1335	1.15	1.48
	115522	AA331393	Hs.47378	ESTs	0.5	3.29
25	115536	AA347193	Hs.62180	ESTs	1	1
	115597	AA411502	Hs.63325	Homo sapiens type II membrane serine pro	1	6.53
	115909	AA436666	Hs.59761	ESTs	1	6.98
	115978	AA447522	Hs.69517	Differentially expressed in Fanconi anem	1	2.31
	116028	AA452112	Hs.42644	thioredoxin-like	0.99	1.68
30	116107	AA456968	Hs.92030	ESTs	1.14	1.8
	116134	AA460246	Hs.50441	CGI-04 protein	1.11	1.86
	116157	AA461063	Hs.44298	Hypothetical protein	0.99	1.9
	116158	AA461187	Hs.61762	Hypoxia-inducible protein 2	0.44	0.86
35	116335	AA495830	Hs.87013	*Homo sapiens cDNA FLJ10238 fis, clone H	0.62	3.89
	116483	C14092	Hs.76118	Ubiquitin carboxyl-terminal esterase L1	1.04	2.36
	117320	N23239	Hs.211092	LUNX protein; PLUNC (palate lung & nasal	0.51	0.64
	117557	N33920	Hs.44532	Diubiquitin	1.11	2.63
	117693	N40939	Hs.112110	PTD007 protein	0.98	1.79
40	117881	N50073	Hs.260622	Butyrate-induced transcript 1	1	1.43
	118368	N64339	Hs.48956	ESTs	0.67	2.86
	118566	N68558	Hs.42824	Hypothetical protein FLJ10718	1.21	0.83
	118695	N71781	Hs.50081	KIAA1199 see CVA7.doc	0.88	1.63
	119780	W72967	Hs.191381	ESTs; Weakly similar to hypothetical pro	1	1
45	119845	W79920	Hs.58561	G protein-coupled receptor 87	1	1
	120102	W95428	Hs.132927	*ESTs, Moderately similar to p53 regulat	1	1
	120104	W95477	Hs.180479	ESTs	0.69	3.07
	120486	AA253400	Hs.137569	Tumor protein 63 kDa with strong homolog	1.08	12.05
	120859	AA350158	Hs.1619	Achaete-scute complex (Drosophila) homol	1	1
50	120880	AA360240	Hs.97019	EST	1	1
	120948	AA397822	Hs.104650	Hypothetical protein FLJ10292	1.04	2.15
	120983	AA398209	Hs.97587	EST	1	1
	121362	AA405500	Hs.97932	Chondromodulin I precursor	1	1
	121369	AA405657	Hs.128791	CGI-09 protein	1	1.8
55	121791	AA423978	Hs.293317	*ESTs, Weakly similar to JM27 [H.sapiens	1	1
	123005	AA479726	Hs.105577	ESTs	1	1
	123044	AA481549	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro	0.95	1.88
	123160	AA488687	Hs.284235	ESTs	1.59	4.98
	123479	AA599469	Hs.135056	clone RP5-850E9 on chromosome 20	1.19	1.64
60	123571	AA608955	Hs.112619	*ESTs, Weakly similar to PQ0109 Purkinje	1.03	1.14
	123829	AA620597	Hs.112208	XAGE-1 protein	1.39	2.2
	124006	D60302	Hs.108977	ESTs	1	4.85
	124059	F13673	Hs.99769	ESTs	1.49	8.62
	124960	T15386	Hs.194766	Seizure related gene 6 (mouse)-like	0.76	0.77
65	125218	W73561	Hs.110024	NADH:ubiquinone oxidoreductase MLRQ subu	1.33	1.77
	125453	R05041	Hs.18048	*Melanoma antigen, family A, 10"	0.8	1.42
	125759	AA425587	Hs.82226	Glycoprotein (transmembrane) nmb	1.52	2.26
	125972	AA434562	Hs.35406	*ESTs, Highly similar to unnamed protein	1.05	2.48
	125994	H55782	Hs.270799	EST	1	1.95
	126395	N70192	Hs.278956	Hypothetical protein FLJ12929	1	1.35
70	126645	AI167942	Hs.61635	STEAP1 (Homo sapiens BAC clone RG041D11	1	2.23
	127221	AI354332	Hs.72365	ESTs	0.73	3.27
	127479	AA513722	Hs.179729	collagen; type X; alpha 1 (Schmid metaph	0.51	1.94
	128192	AI204246		KIAA1085 protein	1.8	3.16
75	128610	L38608	Hs.10247	activated leucocyte cell adhesion molecu	0.89	0.97
	128777	U46006	Hs.10526	Cysteine and glycine-rich protein 2	1	1
	128924	AA234962	Hs.26557	Plakophilin 3	1.3	2.97
	129041	H58873	Hs.169902	*Solute carrier family 2 (facilitated gl	0.84	2.04
	129099	H50398	Hs.108660	*ATP-binding cassette, sub-family C (CFT	0.87	1.04
80	129404	AA172056	Hs.111128	ESTs	1	1
	129466	L42583		*Genbank Homo sapiens keratin 6 isoform	0.72	12.67
	129605	S72493	Hs.115947	Keratin 16 (focal non-epidermolytic palm	0.92	1.5
	129628	U25727	Hs.1174	*Cyclin-dependant kinase inhibitor 2A (m	0.85	1.93
	130023	X13461	Hs.239600	Calmodulin-like 3	0.84	1.22
	130080	X14850	Hs.147097	*H2A histone family, member X"	0.98	1.96
85	130385	AA126474	Hs.155223	stanniocalcin 2	1	1

	130410	V01514	Hs.155421	Alpha-fetoprotein	0.63	0.63
	130441	U35835	Hs.301387	"Human DNA-PK mRNA, partial cds"	1.15	3.65
	130482	L32866	Hs.1578	Baculoviral IAP repeat-containing 5 (sur	1	1.88
5	130553	AA430032	Hs.252587	Pituitary tumor-transforming 1	0.92	1.95
	130577	M35410	Hs.162	Insulin-like growth factor binding prote	1.17	4.7
	130527	L23303	Hs.1695	Matrix metalloproteinase 12 (macrophage	0.69	4.05
	130800	AA223385	Hs.19574	ESTs; Weakly similar to katanin p80 subu	1.13	2.41
	130939	AA598663	Hs.21400	ESTs	0.8	0.89
10	131046	X02530	Hs.2248	INTERFERON-GAMMA INDUCED PROTEIN PRECURS	0.8	1.15
	131244	D38076	Hs.24763	RAN binding protein 1	1.13	1.85
	131877	J04088	Hs.156346	Topoisomerase (DNA) II alpha (170kD)	1	1
	131927	AA461549	Hs.34780	"Doublecortin; lissencephaly, X-linked (0.81	0.62
	131965	W90146	Hs.35962	ESTs	0.74	3.27
15	131978	D80008	Hs.36232	KIAA0186 gene product	1	1
	132354	L05187	Hs.211913	Small proline-rich protein 1A	0.69	1.43
	132543	AA417152	Hs.5101	ESTs; Highly similar to protein regulati	0.79	4.27
	132632	N59764	Hs.5398	guanine-monophosphate synthetase	1	1.08
	132653	U31201	Hs.54451	"laminin gamma2 chain gene (LINC2), exon	1	1
20	132659	Z75190	Hs.54481	"Low density lipoprotein receptor-relate	0.89	0.89
	132710	W93726	Hs.55279	"Serine (or cysteine) proteinase inhibit	0.64	4.41
	132758	W52432	Hs.55105	"ESTs, Weakly similar to WDNM RAT WDNM1	1.55	2.08
	132767	L05188	Hs.231622	Small proline-rich protein 2B	0.83	1.66
	132816	M74542	Hs.575	Aldehyde dehydrogenase 3	0.55	0.55
25	132990	AA458761	Hs.18387	transcription factor AP-2 alpha (activat	1	3.53
	133070	U69511	Hs.64311	"A disintegrin and metalloproteinase dom	1.16	2
	133282	U52950	Hs.286145	"SRB7 (suppressor of RNA polymerase B, y	1	2.7
	133317	AA215239	Hs.70830	U6 snRNA-associated Sm-like protein LSM7	0.95	1.42
	133370	AA156897	Hs.72157	Homo sapiens mRNA; cDNA DKFZp564I1922	1.12	2.55
30	133391	X57579	Hs.727	H.sapiens activin beta-A subunit (exon 2	1.65	1.76
	133832	H03387	Hs.241305	estrogen-responsive B box protein (EBBP)	1.02	1.39
	134032	Z81326	Hs.78589	"Serine (or cysteine) proteinase inhibit	1	1
	134168	AA398908	Hs.181634	"Homo sapiens cDNA: FLJ23502 fis, clone	0.95	1.53
	134218	AA227480	Hs.80205	Pim-2 oncogene	1.36	2.48
35	134405	R67275	Hs.82772	"collagen, type XI, alpha 1"	0.76	2.86
	134453	X70683	Hs.83484	SRY (sex determining region Y)-box 4	1.89	3.78
	134470	X54942	Hs.83758	CDC28 protein kinase 2	1.82	4.11
	134645	U87459	Hs.167379	"Cancer/hastis antigen (NY-ESO-1, CTAG1,	0.82	0.83
	134781	M17183	Hs.89626	Parathyroid hormone-like hormone	1	1
40	135002	U19147	Hs.272484	G antigen 6	1	1
	100040	M97935		AFFX control: STAT1	0.92	1.25
	101201	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin;	2.92	8.5
	101664	M50752	Hs.121017	H2A histone family; member A	1	1
	102025	U03911	Hs.78934	mult S (E. coli) homolog 2 (colon cancer;	0.8	1.61
45	102031	U04898	Hs.2156	RAR-related orphan receptor A	1	1
	102221	U24576		LIM domain only 4	1	1
	102270	U30255	Hs.75888	phosphogluconate dehydrogenase	1.08	1.43
	102339	U37022	Hs.95577	cyclin-dependent kinase 4	0.88	1.32
	102391	U41668	Hs.77494	deoxyguanosine kinase	1.07	1.58
50	103000	X51956	Hs.146580	enolase 2; (gamma; neuronal)	0.91	1.49
	103395	X94754	Hs.119503	methionine-tRNA synthetase	0.89	1.32
	105638	AA281599	Hs.20418	Homo sapiens mRNA for histone H2B; c	0.91	1.25
	105726	AA292328	Hs.9754	activating transcription factor 5	0.94	1.48
	114841	AA234722	Hs.55408	ESTs; Moderately similar to CALCIUM-DEPE	0.78	1.56
55	115206	AA262491	Hs.186572	ESTs	1	1
	115906	AA436616	Hs.82302	ESTs	0.74	2.52
	119132	R49046	Hs.107911	ATP-binding cassette; sub-family B (MDR/	1.1	1.51
	124163	H30539	Hs.189838	ESTs	1	1
	126487	AA482505	Hs.184601	solute carrier family 7 (cationic amino	1.01	1.46
60	127141	AA307950	Hs.75478	KIAA0956 protein	0.85	1.4
	128034	AA905754	Hs.75103	tyrosine 3-monooxygenase/tryptophan 5-mo	1	1.18
	128609	AA234365	Hs.102456	survival of motor neuron protein interac	1	1.5
	128895	R37753	Hs.106985	ESTs	1.7	2
	130199	Z48579	Hs.172028	a disintegrin and metalloprotease domain	1	1
65	130524	U89995	Hs.159234	forkhead box E1	1	1
	133000	U24152	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast	1	1
	133658	M25756	Hs.75426	secretogranin II (chromogranin C)	1	1
	135047	AA460466	Hs.93597	ESTs	1	1
70	100053	M27830		AFFX control: 28S ribosomal RNA	0.88	1.53
	100114	D00596	Hs.82962	thymidylate synthetase	0.68	1.86
	100128	D11094	Hs.61153	proteasome (prosome; macropain) 26S subu	1.29	2.03
	100154	D14657	Hs.81892	KIAA0101 gene product	0.71	4.26
	100161	D14694	Hs.77329	phosphatidylserine synthase 1	1.02	1.56
	100168	D14874	Hs.394	adrenomedullin	0.46	1.17
75	100187	D17793	Hs.78183	aldo-keto reductase family 1; member C3	1	1
	100188	D21063	Hs.57101	minichromosome maintenance deficient (S.	0.97	1.4
	100217	D26600	Hs.89545	proteasome (prosome; macropain) subunit;	1.13	1.9
	100220	D28364		"Human mRNA for annexin II, 5'UTR (seq	1.11	1.53
	100287	D43950	Hs.1600	chaperonin containing TCP1; subunit 5 (e	1.13	2.09
	100297	D49489	Hs.182429	protein disulfide isomerase-related prot	0.92	1.78
80	100330	D55716	Hs.77152	minichromosome maintenance deficient (S.	1.07	1.61
	100355	D78129		"Homo sapiens mRNA for squalene epoxid	0.96	1.87
	100364	D78585	Hs.154858	carbamoyl-phosphate synthetase 2; aspart	1.49	2.46
	100368	D79987	Hs.153479	extra spindle poles; S. cerevisiae; homo	0.59	1.32
	100398	D84557	Hs.155462	minichromosome maintenance deficient (mi	1.08	1.9
85	100438	D87448	Hs.91417	topoisomerase (DNA) II binding protein	1	2.15

	100455	D87953	Hs.75789	N-myc downstream regulated	0.91	1.48
	100491	HG1153-HT1153		Nucleoside Diphosphate Kinase Nm23-H2s	0.99	1.41
	100518	HG174-HT174		Desmoplakin I	1.28	3.17
	100528	HG1828-HT1857		***Nexin, Glia-Derived***	0.68	1.9
5	100661	HG2874-HT3018		Ribosomal Protein L39 Homolog	1.1	5.44
	100667	HG2981-HT3127		***Epican, Alt. Splice 11***	0.8	1.97
	100830	HG4074-HT4344		Rad2	1.01	2.12
	101061	K03515	Hs.944	glucose phosphate isomerase	0.91	1.79
	101131	L10838	Hs.167460	splicing factor; arginine/serine-rich 3	1.23	1.87
10	101162	L14595	Hs.174203	solute carrier family 1 (glutamate/neutral)	1.35	2.73
	101181	L19586	Hs.73798	macrophage migration inhibitory factor (1.03	1.78
	101183	L19779	Hs.795	H2A histone family; member O	0.57	1.3
	101216	L25876	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	0.7	2.2
	101228	L27705	Hs.82916	chaperonin containing TCP1; subunit 6A (0.99	1.99
15	101233	L29008	Hs.878	sorbitol dehydrogenase	0.82	2.11
	101247	L33801	Hs.78802	glycogen synthase kinase 3 beta	1.2	1.91
	101332	L47276		***Homo sapiens (cell line HL-6) alpha t	0.69	2.78
	101342	L76191	Hs.182018	interleukin-1 receptor-associated kinase	1.04	1.84
	101396	M15796	Hs.78995	proliferating cell nuclear antigen	0.95	3.55
20	101423	M18391	Hs.89839	EphA1	1	1.5
	101445	M21259	Hs.1066	small nuclear ribonucleoprotein polypept	1.21	1.96
	101505	M27395	Hs.75692	asparagine synthetase	0.93	1.6
	101525	M29535	Hs.12163	eukaryotic translation initiation factor	1.19	1.93
	101535	M30448	Hs.251669	casein kinase 2; beta polypeptide	0.96	1.42
25	101607	M38690	Hs.1244	CD9 antigen (p24)	1.11	1.25
	101624	M55998		***Human alpha-1 collagen type I gene, 3	1.17	1.98
	101758	M77835	Hs.79217	pyrroline-5-carboxylate reductase 1	1.77	3.45
	101839	M93035	Hs.692	membrane component; chromosomal 4; surfa	0.71	1.45
	101853	M94362	Hs.76084	lamin B2	0.84	1.19
30	101977	S83364		***putative Rab5-interacting protein (cl	0.89	1.9
	101992	U01038	Hs.77597	polo (Drosophila)-like kinase	0.66	1.46
	102009	U02680	Hs.82643	protein tyrosine kinase 9	1.23	3.35
	102012	U03057	Hs.118400	singed (Drosophila)-like (sea urchin fas	0.85	1.88
35	102039	U05861	Hs.201957	aldo-keto reductase family 1; member C1	0.93	2.32
	102123	U14518	Hs.1594	centromere protein A (17kD)	1	4.28
	102130	U15009	Hs.1575	small nuclear ribonucleoprotein D3 polyp-	0.89	1.42
	102148	U16954	Hs.75823	ALL1-fused gene from chromosome 1q	0.8	2.95
	102210	U23028	Hs.2437	eukaryotic translation initiation factor	1.01	1.34
	102220	U24389	Hs.65436	lysyl oxidase-like 1	1.15	2.34
40	102260	U28386	Hs.159557	karyopherin alpha 2 (RAG cohort 1; impor	1.14	2.69
	102330	U35451	Hs.77254	chromobox homolog 1 (Drosophila HP1 beta	1.05	1.7
	102423	U44754	Hs.179312	small nuclear RNA activating complex; po	1.14	2.99
	102455	U48705	Hs.75552	discoidin domain receptor family; member	1.05	2.01
45	102499	U51478	Hs.76941	ATPase; Na+/K+ transporting; beta 3 poly	1.27	1.92
	102522	U53347	Hs.183556	solute carrier family 1 (neutral amino a	0.84	1.31
	102590	U62136		***Homo sapiens enterocyte differentiati	1.11	1.6
	102676	U72514	Hs.12045	putative protein	1.04	2.17
	102687	U73379	Hs.93002	ubiquitin carrier protein E2-C	0.86	2.28
50	102704	U76638	Hs.54089	BRCA1 associated RING domain 1	1.12	1.63
	102781	U83843		***Human HIV-1 Nef interacting protein (0.9	1.39
	102784	U85658	Hs.61796	transcription factor AP-2 gamma (activat	0.98	2.16
	102827	U91327	Hs.6456	chaperonin containing TCP1; subunit 2 (b	0.96	1.62
	102935	X13482	Hs.80506	small nuclear ribonucleoprotein polypept	1.21	4.2
55	102972	X16652	Hs.87268	annexin A8	1.25	2.32
	102983	X17620	Hs.118638	non-metastatic cells 1; protein (NM23A)	1.03	1.83
	103023	X53793	Hs.117950	multifunctional polypeptide similar to S	1.58	5.44
	103038	X54941	Hs.77550	CDG28 protein kinase 1	1.32	3.79
	103075	X59543	Hs.2934	ribonucleotide reductase M1 polypeptide	1.11	2.68
60	103168	X68314	Hs.2704	glutathione peroxidase 2 (gastrointestinal	0.75	3.05
	103185	X69910	Hs.74368	transmembrane protein (63kD); endoplasmic	1.01	1.97
	103212	X73874	Hs.2393	phosphorylase kinase; alpha 1 (muscle)	0.95	1.72
	103223	X74801	Hs.1708	chaperonin containing TCP1; subunit 3 (g	0.97	1.77
	103260	X78416	Hs.3155	casein; alpha	1	1
65	103262	X78565	Hs.204133	hexabrachion (lamin C; cytactin)	1.23	3.09
	103330	X85373	Hs.77496	small nuclear ribonucleoprotein polypept	1.12	2.25
	103364	X90872	Hs.75854	SULT1C sulfotransferase	2.85	4.62
	103375	X91868	Hs.54416	sine oculis homeobox (Drosophila) homolo	1	2.48
	103391	X94453	Hs.114366	pyrroline-5-carboxylate synthetase (glut	1	1.53
70	103404	X95586	Hs.78596	proteasome (prosome; macropain) subunit	0.92	1.53
	103437	X98260	Hs.82254	M-phase phosphoprotein 11	0.92	1.54
	103448	X99133	Hs.204238	lipocatin 2 (oncogene 24p3)	0.55	0.96
	103505	Z35402	Hs.194657	cadherin 1; E-cadherin (epithelial)	1.32	2.51
	103646	Z68228	Hs.2340	junction plakoglobin	0.88	1.28
75	103658	Z74615	Hs.172928	collagen; type I; alpha 1	1.06	2.98
	103774	AA092898	Hs.92918	ESTs; Weakly similar to R07G3.8 (C.elega	1.88	4.66
	104261	AF008442	Hs.5409	RNA polymerase I subunit	0.87	2.47
	104276	C02193	Hs.85222	ESTs; Weakly similar to R27090_2 (H.sapi	1.4	2.49
	104289	C16281	Hs.75478	KIAA0955 protein	1.15	1.68
	104434	L02870	Hs.1640	collagen; type VII; alpha 1 (epidermolys	1.04	1.49
80	104453	M19169	Hs.123114	cystatin SN	0.38	0.76
	104611	R38280	Hs.125845	ribulose-5-phosphate-3-epimerase	1.08	2.25
	104758	AA024661	Hs.7010	ESTs; Weakly similar to ACYL-COA DEHYDRO	1.14	1.65
	105114	AA156532	Hs.11801	adenosine A2b receptor pseudogene	0.91	1.39
	105132	AA159501	Hs.247280	HBV associated factor	1.08	1.7
85	105174	AA186613	Hs.34744	ESTs	0.95	2.05

	105280	AA232215	Hs.14600	ESTs	1	1.4
	105344	AA235303	Hs.8645	ESTs	0.72	2.02
	105516	AA257971	Hs.21214	ESTs	1.35	3.56
5	105521	AA280865	Hs.6375	Homo sapiens mRNA; cDNA DKFZp564K0222 (f	1.23	1.82
	105698	AA287393	Hs.15202	ESTs; Weakly similar to oligodendrocyte	0.98	1.28
	105705	AA290767	Hs.101282	Homo sapiens mRNA; cDNA DKFZp434B102 (fr	0.92	1.32
	105724	AA292098	Hs.22934	ESTs; Weakly similar to ZINC FINGER PROT	0.99	1.41
	105782	AA350215	Hs.21580	ESTs	1	1
	105799	AA372018	Hs.24743	ESTs	1.08	1.78
10	105807	AA393803	Hs.16869	ESTs; Moderately similar to COLLAGEN ALP	0.95	1.34
	105891	AA400768	Hs.26662	ESTs; Weakly similar to tumor necrosis f	0.87	2.25
	105936	AA404338		ESTs	1.14	1.46
	106069	AA417741	Hs.29899	ESTs; Weakly similar to ZINC FINGER PROT	1	1
	106103	AA421104	Hs.12094	ESTs	1.04	1.44
15	106140	AA424524	Hs.14912	KIAA0266 protein	1.23	2.11
	106149	AA424881	Hs.256301	ESTs	0.83	1.48
	106154	AA425304	Hs.6994	ESTs	0.77	2.05
	106182	AA426609	Hs.10862	ESTs	0.74	2.23
	106220	AA426582	Hs.32196	ESTs; Moderately similar to metargidin p	0.97	1.99
20	106228	AA429290	Hs.17719	ESTs	0.99	1.54
	106318	AA436570	Hs.9605	pre-mRNA cleavage factor Irm (25kD)	0.95	2.09
	106341	AA441798	Hs.5243	ESTs; Moderately similar to pIL2 hypothe	0.98	2.66
	106432	AA448850	Hs.17138	ESTs	0.95	1.93
	106474	AA450212	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (fr	1	1
25	106483	AA451676	Hs.30299	IGF-II mRNA-binding protein 2	1.4	2.29
	106599	AA457235	Hs.12842	ESTs; Moderately similar to non-function	1	1.82
	106611	AA458904	Hs.26267	ESTs; Weakly similar to torsinA [H.sapie	1.49	2.78
	106654	AA460449	Hs.3784	ESTs; Highly similar to phosphoserine am	1	1.4
	107076	AA609145	Hs.21143	ESTs; Weakly similar to fos39554_1 [H.s	1.11	1.49
30	107115	AA610108	Hs.27693	ESTs; Highly similar to CGL-124 protein	1	1.03
	107129	AA620553	Hs.4756	flap structure-specific endonuclease 1	1.13	3.63
	107159	AA621340	Hs.10600	ESTs; Weakly similar to ORF YKR081c [S.c	1.05	2.09
	107444	W28391	Hs.5181	proliferation-associated 2G4; 38kD	1.18	1.9
	107481	W58247	Hs.27437	Homo sapiens kinesin superfamily motor K	0.99	2.74
35	107516	X56597	Hs.99853	fibrillarin	0.94	1.77
	107529	Y12065	Hs.5092	nucleolar protein (KKE/D repeat)	1.05	2.29
	107531	Y13936	Hs.17883	protein phosphatase 1G (formerly 2C); ma	1.06	1.62
	107801	AA019433	Hs.173100	ESTs	1.03	1.4
	107957	AA031948	Hs.57548	ESTs	0.95	1.46
40	108565	AA085342	Hs.1526	ATPase; Ca++ transporting; cardiac muscl	0.59	1.35
	108780	AA128581	Hs.117938	collagen; type XVII; alpha 1	1	7.63
	108828	AA131584	Hs.71435	DKFZP564O0463 protein	1.33	2.56
	109060	AA160879	Hs.241551	chloride channel; calcium activated; fam	0.67	1.42
	109112	AA169379	Hs.72865	ESTs	1.03	2.31
45	109344	AA213696	Hs.86559	poly(A)-binding protein-like 1	0.97	1.55
	109412	AA227145	Hs.209473	ESTs; Weakly similar to REGULATOR OF MIT	0.76	1.87
	110780	N23174	Hs.22891	solute carrier family 7 (cationic amino	0.9	0.95
	110958	N50550	Hs.24587	signal transduction protein (SH3 contain	1.17	2.26
	111018	N54067	Hs.3828	mitogen-activated protein kinase kinase	1.21	1.85
50	111337	N79512	Hs.16607	ESTs; Highly similar to Myosin heavy cha	1	1.45
	112305	R54822	Hs.26244	ESTs	1	1
	112401	R61279	Hs.237536	ESTs; Weakly similar to F25B5.3 [C.elega	1.24	1.64
	112853	T02843	Hs.4351	EST	1.56	1.96
	112869	T03313	Hs.4747	dyskeratosis congenita 1; dyskerin	1.03	1.57
55	112992	T23513	Hs.7147	ESTs	1	1
	113048	T25895	Hs.184008	ESTs; Weakly similar to RNA-binding prot	1.37	2.26
	113063	T32438	Hs.5027	ESTs	1	1
	113179	T55182	Hs.152571	ESTs; Highly similar to IGF-II mRNA-bind	1.33	2.7
60	113573	T91166	Hs.15990	ESTs	0.76	1.47
	113811	W44928	Hs.4878	ESTs	0.79	1.51
	114086	Z38266	Hs.12770	Homo sapiens PAC clone DJ0777023 from 7p	0.9	1.34
	114587	AA070827	Hs.180320	ESTs; Weakly similar to GOLGI 4-TRANSMEM	1.02	1.76
	114846	AA234929	Hs.44343	ESTs	1.32	2.36
65	114964	AA243873	Hs.82184	ring finger protein 3	1.1	1.84
	115047	AA252627	Hs.22554	homeo box B5	1.01	2.36
	115166	AA258409	Hs.196907	myelin protein zero-like 1	1.05	2.31
	115167	AA258421	Hs.43728	hypothetical protein	1.52	2.52
	115239	AA278650	Hs.73291	ESTs; Weakly similar to similar to the b	0.7	2.57
70	115278	AA279757	Hs.67466	ESTs; Weakly similar to BACN32G11.d [D.m	1.14	2.12
	115652	AA405098	Hs.39178	ESTs	0.82	4.67
	115875	AA433943	Hs.43946	ESTs; Weakly similar to Weak similarity	1.2	1.98
	116004	AA449122	Hs.76086	ESTs; Highly similar to small zinc finge	0.96	1.31
	116121	AA459254	Hs.48855	ESTs	0.97	1.55
	116129	AA459956	Hs.49163	ESTs; Highly similar to putative ribonuc	1.08	2.73
75	116190	AA464963	Hs.67776	ESTs	0.8	1.57
	116312	AA490494	Hs.65403	ESTs	1.37	2.65
	116732	F13779	Hs.165909	ESTs	0.92	1.8
	117602	N35020	Hs.44685	ESTs; Weakly similar to GOLIATH PROTEIN	1.15	1.84
80	117950	N51394	Hs.75478	KIAA0956 protein	1.04	2.36
	117992	N52000	Hs.172089	Homo sapiens mRNA; cDNA DKFZp586B0222 (f	0.62	1.29
	118785	N75386	Hs.111867	GLI-Kruppel family member GLI2	1	1
	119717	W69134	Hs.57937	ESTs	1	1.4
	119814	W74069	Hs.58350	ESTs	0.78	1.77
	120128	Z38499	Hs.91448	MKP-1 like protein tyrosine phosphatase	0.86	1.46
85	120242	Z98443	Hs.85366	ESTs	0.83	2.01

	120483	AA252994	Hs.1578	apoptosis inhibitor 4 (survivin)	0.74	1.64
	121054	AA398604	Hs.97387	ESTs	1.05	1.93
	121326	AA404246	Hs.97031	ESTs; Weakly similar to Similar to phyto	0.98	1.3
	121376	AA405699	Hs.166232	ESTs; Moderately similar to SODIUM- AND	0.91	1.83
5	121457	AA411448	Hs.208985	ESTs	0.91	1.59
	121780	AA422086	Hs.124660	ESTs	0.46	0.55
	121781	AA422150	Hs.98370	cytochrome P540 family member predicted	1.07	1.54
	121844	AA425732	Hs.98485	gap junction protein; beta 2; 26kD (conn	0.94	1.4
	122059	AA431737	Hs.93749	EST	1.93	2.33
10	122338	AA443311	Hs.98998	ESTs	1	1
	122354	AA443772	Hs.196692	ESTs	0.88	1.39
	122591	AA453265	Hs.99311	ESTs; Weakly similar to MRJ [H.sapiens]	2.28	2.93
	122790	AA460156	Hs.99556	ESTs	0.88	1.3
	123398	AA521265	Hs.105514	ESTs	1	1.93
15	123518	AA608531	Hs.170313	ESTs	1	1
	123673	AA609471	Hs.112712	ESTs	1	1.15
	124000	D57317	Hs.74861	activated RNA polymerase II transcriptio	0.74	1.12
	124367	N24006	Hs.99348	distal-less homeo box 5	0.67	1.1
	124447	N48000	Hs.140945	Homo sapiens mRNA; cDNA DKFZp586L141 (fr	1.19	1.7
20	125756	W25498	Hs.81634	ATP synthase; H+ transporting; mitochond	0.93	1.59
	125769	AA382972	Hs.82128	ST4 oncofetal trophoblast glycoprotein	1.65	6.76
	125852	H09290	Hs.76550	Homo sapiens mRNA; cDNA DKFZp564B1264 (f	0.72	2.26
	125924	AA526849	Hs.82109	syndecan 1	1.22	2.25
	126037	M85772	Hs.6066	KIAA1112 protein	1.36	1.63
25	126214	N29455	Hs.74316	desmoplakin (DP; DPL)	1.93	3.55
	126414	N78770	Hs.223439	ESTs	1.21	1.66
	126737	AA488132	Hs.62741	ESTs	1	1
	126743	AA179253	Hs.172182	poly(A)-binding protein; cytoplasmic 1	1.3	2.16
	126926	AA179546	Hs.832	ESTs; Highly similar to INTEGRIN BETA-8	2.53	2.8
30	127432	AA501734	Hs.170311	heterogeneous nuclear ribonucleoprotein	1.57	2.12
	128218	H02682	Hs.99189	ESTs; Moderately similar to recombinatio	1.24	2.09
	128527	M31523	Hs.101047	transcription factor 3 (E2A immunoglobul	1.08	1.78
	128568	X60573	Hs.247568	adenylate kinase 3	1.23	3.48
	128584	M11433	Hs.101850	retinol-binding protein 1; cellular	0.87	2.42
35	128628	C14037	Hs.251978	EST	1.22	1.9
	128691	W27939	Hs.103834	ESTs	1.1	1.73
	128714	V00599	Hs.179661	Homo sapiens clone 24703 beta-tubulin mR	0.92	1.17
	128733	AA328993	Hs.104558	ESTs	1.34	1.94
40	128781	X85372	Hs.105465	small nuclear ribonucleoprotein polypept	0.9	1.34
	129052	AA496297	Hs.182740	ribosomal protein S11	2.59	3.19
	129095	L12350	Hs.108623	thrombospondin 2	1.04	3.2
	129241	AA435665	Hs.109706	ESTs; Moderately similar to HN1 (M.muscu	0.95	1.61
	129665	M88458	Hs.118778	KDEL (Lys-Asp-Glu-Leu) endoplasmic retic	1.28	2.63
	129703	AA401348	Hs.179999	ESTs	0.97	1.63
45	129720	AA476582	Hs.12152	ESTs; Moderately similar to SIGNAL RECOG	1.09	1.79
	129850	N20593	Hs.56845	GDP dissociation inhibitor 2	0.74	1.68
	129896	AA043021	Hs.13225	UDP-Gal:betaGlcNAc beta 1;4- galactosylt	1.43	4.19
	130069	AA055896	Hs.146428	collagen; type V; alpha 1	1.17	1.98
50	130405	H88359	Hs.155396	nuclear factor (erythroid-derived 2)-lik	1.26	1.79
	130541	X05608	Hs.211584	neurofilament; light polypeptide (68kD)	1	1
	130599	M91670	Hs.174070	ubiquitin carrier protein	1.07	1.66
	130867	J04093	Hs.2056	UDP glycosyltransferase 1	1	4.8
	131009	AA063596	Hs.22142	ESTs; Weakly similar to NADH-CYTOCHROME	0.93	1.05
55	131028	U20240	Hs.2227	CCAAT/enhancer binding protein (C/EBP);	1	1.23
	131083	U66661	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	1.1	1.8
	131091	T35341	Hs.22880	ESTs; Highly similar to dipeptidyl pepti	1.28	1.98
	131144	C14412	Hs.23528	ESTs; Highly similar to HSPC038 protein	1.43	2.06
	131148	C00038	Hs.23579	ESTs	0.88	3.38
	131164	Y00503	Hs.182265	keratin 19	1.19	2.77
60	131185	M25753	Hs.23960	cyclin B1	0.86	3.84
	131219	C00476	Hs.24395	small inducible cytokine subfamily B (Cy	0.66	2.96
	131454	AA455896	Hs.2699	glyptcan 1	0.99	1.54
	131587	L11066	Hs.3069	heat shock 70kD protein 9B (mortalin-2)	1	1.18
	131689	AA599653	Hs.30696	transcription factor-like 5 (basic helix	1	1.95
65	131692	D50914	Hs.30736	KIAA0124 protein	1.55	2.39
	131786	AA135554	Hs.32125	ESTs	1	1.33
	131843	AA195893	Hs.184062	ESTs; Moderately similar to putative Rab	0.83	1.63
	131860	U02082	Hs.334	Oncogene TIM	1.08	2.2
	131884	H90124	Hs.3453	ribosomal protein S23	1.23	1.24
70	131903	AA481723	Hs.3436	deleted in oral cancer (mouse; homolog)	0.91	1.18
	131945	M87339	Hs.35120	replication factor C (activator 1) 4 (37	1	2.8
	131958	AA093998	Hs.3566	ESTs; Highly similar to phosphorylation	0.87	1.36
	131964	W42508	Hs.3593	ESTs	1	1.25
	132001	J00277	Hs.37003	v-Ha-ras Harvey rat sarcoma viral oncoge	1.12	1.43
75	132040	AA146843	Hs.172894	BH3 interacting domain death agonist	1	1.55
	132055	D82226	Hs.211594	proteasome (prosome; macropain) 26S subu	0.89	1.27
	132109	AA599801	Hs.40098	ESTs	1	1.05
	132112	AA150661	Hs.40154	junonji (mouse) homolog	0.99	1.44
	132123	AA447123	Hs.250705	ESTs	1.06	2.46
80	132162	H89551	Hs.41241	ESTs	1.08	2.46
	132180	AA405569	Hs.418	fibroblast activation protein; alpha; se	1.02	4.56
	132309	AA460917	Hs.2780	jun D proto-oncogene	1.16	1.8
	132371	AA235448	Hs.46677	ESTs	0.8	1.26
	132618	AA253330	Hs.5344	adaptor-related protein complex 1; gamma	0.5	1.49
85	132736	U68019	Hs.211578	MAD (mothers against decapentaplegic; Dr	1.21	1.81

	132771	AA488432	Hs.56407	phosphoserine phosphatase	1	1.3
	132833	U78525	Hs.57783	eukaryotic translation initiation factor	0.91	1.43
	132922	T23641	Hs.6066	KIAA1112 protein	1.16	1.53
5	132959	AA028103	Hs.61472	ESTs; Weakly similar to unknown JS.cerev	1.02	1.88
	132994	AA505133	Hs.7594	solute carrier family 2 (facilitated glu	0.72	2.97
	133005	C21400	Hs.103329	KIAA0970 protein	0.88	1.34
	133065	X62535	Hs.172690	diacylglycerol kinase; alpha (80kD)	0.93	1.23
	133083	N70633	Hs.6456	chaperonin containing TCP1; subunit 2 (b	1.14	1.76
10	133086	L17131	Hs.133800	high-mobility group (nonhistone chromoso	0.97	1.43
	133134	T89703	Hs.65648	RNA binding motif protein 8	1.1	1.8
	133195	AA350744	Hs.181409	KIAA1007 protein	2.29	2.69
	133313	AA249427	Hs.70704	ESTs	1.07	1.68
	133331	T62039	Hs.158675	ribosomal protein L14	0.85	1.18
15	133438	D13370	Hs.73722	APEX nuclease (multifunctional DNA repai	0.91	1.45
	133445	T99303	Hs.73797	guanine nucleotide binding protein (G pr	0.94	1.68
	133483	X52426	Hs.74070	keratin 13	0.85	1.14
	133492	L40397	Hs.74137	transmembrane trafficking protein	1.1	1.69
	133504	W95070	Hs.74316	desmoplakin (DPI; DPII)	0.7	6.21
	133517	X52947	Hs.74471	gap junction protein; alpha 1; 43kD (con	0.95	1.3
20	133540	D78151	Hs.74619	proteasome (prosome; macropain) 26S subu	0.91	1.25
	133594	L07758	Hs.172589	nuclear phosphoprotein similar to S. cer	0.84	1.29
	133627	U09587	Hs.75280	glycyl-tRNA synthetase	1.09	1.99
	133671	T25747	Hs.75471	zinc finger protein 146	1.02	1.5
	133859	U86782	Hs.178761	26S proteasome-associated pad1 homolog	1.11	3.33
25	133865	F09315	Hs.170290	discs; large (Drosophila) homolog 5	1.84	6.7
	133913	W84712	Hs.7753	calumenin	1.15	1.86
	133963	L34587	Hs.184693	transcription elongation factor B (SII)	1.3	1.91
	133982	U47621	Hs.207251	nucleolar autoantigen (55kD) similar to	1.3	1.99
	134100	L07540	Hs.171075	replication factor C (activator 1) 5 (36	0.72	1.65
30	134110	U41060	Hs.79135	LIV-1 protein; estrogen regulated	1.04	1.62
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	1	1.55
	134161	U97188	Hs.79440	IGF-II mRNA-binding protein 3	0.82	1.95
	134193	F09570	Hs.7980	ESTs	0.98	1.48
35	134367	X54199	Hs.82285	phosphoribosylglycinamide formyltransfer	1	2.8
	134402	U25165	Hs.82712	fragile X mental retardation; autosomal	1.26	2
	134457	D86963	Hs.174044	dishevelled 3 (homologous to Drosophila	1	1.47
	134469	X17567	Hs.83753	small nuclear ribonucleoprotein polypept	0.94	1.57
	134498	M63180	Hs.84131	threonyl-tRNA synthetase	1.2	2.64
40	134501	W84870	Hs.211558	eukaryotic translation initiation factor	0.84	1.36
	134507	M63488	Hs.84318	replication protein A1 (70kD)	1.7	2.93
	134548	U41515	Hs.85215	Deleted in split-hand/split-foot 1 regio	1.46	2.73
	134599	X99226	Hs.86297	Fanconi anemia; complementation group A	1.36	2.22
	134692	R73567	Hs.8850	a disintegrin and metalloproteinase doma	0.77	1.64
45	134693	N70361	Hs.8854	ESTs	1.09	1.82
	134806	Z49099	Hs.89718	spermine synthase	0.98	1.35
	134821	Z34974	Hs.198382	plakophilin 1 (ectodermal dysplasia/skin	0.99	1.4
	134864	Y08999	Hs.90370	actin related protein 2/3 complex; subun	0.95	1.42
	134914	U29615	Hs.91093	chitinase 1 (chitinohydrolase)	1.18	1.29
50	134953	L10678	Hs.91747	profilin 2	0.95	1.76
	134993	AA282343	Hs.9242	purine-rich element binding protein B	0.98	1.73
	135051	C15324	Hs.93668	ESTs	1.35	2.11
	135158	U51711		Human desmocollin-2 mRNA; 3' UTR	0.86	1.16

Table 1B shows the accession numbers for those pkys in Table 1A lacking unigenelD's. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the Accession column.

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT	Accessions
65	100661	23182_1
	100667	26401_3
70		
75	100668	26401_3
80		
85	101332	25130_1

BE623001 L05095 AA383604 AW966416 N53295 AA460213 AW571519 AA603655
 L05424 X56794 S66400 X55150 W60071 AW351820 X55938 M83326 BE005289 BE070059 M83324 BE005248 BE069717 BE181648 BE069700
 AW606203 BE069721 AW382138 AW803776 BE453954 BE005334 BE005274 T27386 AA932714 AA972695 AW377728 A1632506 T29066
 A1783934 AW377727 BE163715 AL047291 AA279047 AA523003 BE008048 BE440141 W23614 BE090519 BE092193 N29181 N20358 N44153
 BE546944 T69231 AW377441 AA907406 H50799 AW051416 A1420712 BE620922 A1279161 AA992549 W47198 BE005241 A1342696 H50700
 A1969974 A1863855 AA374490 AW130675 A1950633 AA146687 H99482 X55150 BE005414 BE005339 N28294 A1673068 A1887890 AW804171
 A1675961 AW804172 AA778841 AL048050 A1127757 A1095568 AW204965 AW468978 W31898 A1052595 A1278771 BE464018 A1081503 A1824196
 AA513211 AA411062 AW084376 N48752 AA703209 N35580 AW059918 AA054563 A1280942 T27619 BE621435 N66010 AW589527 A1160414
 AA283090 AA962535 H82726 W52115 W45432 W60433 AA577548 AA146714 BE150994 AA054615 AW796025 AW382768 BE555671 C00444
 AA054555
 L05424 X56794 S66400 X55150 W60071 AW351820 X55938 M83326 BE005289 BE070059 M83324 BE005248 BE069717 BE181648 BE069700
 AW606203 BE069721 AW382138 AW803776 BE453954 BE005334 BE005274 T27386 AA932714 AA972695 AW377728 A1632506 T29066
 A1783934 AW377727 BE163715 AL047291 AA279047 AA523003 BE008048 BE440141 W23614 BE090519 BE092193 N29181 N20358 N44153
 BE546944 T69231 AW377441 AA907406 H50799 AW051416 A1420712 BE620922 A1279161 AA992549 W47198 BE005241 A1342696 H50700
 A1969974 A1863855 AA374490 AW130675 A1950633 AA146687 H99482 X55150 BE005414 BE005339 N28294 A1673068 A1887890 AW804171
 A1675961 AW804172 AA778841 AL048050 A1127757 A1095568 AW204965 AW468978 W31898 A1052595 A1278771 BE464018 A1081503 A1824196
 AA513211 AA411062 AW084376 N48752 AA703209 N35580 AW059918 AA054563 A1280942 T27619 BE621435 N66010 AW589527 A1160414
 AA283090 AA962535 H82726 W52115 W45432 W60433 AA577548 AA146714 BE150994 AA054615 AW796025 AW382768 BE555671 C00444
 AA054555
 J04088 NM_001067 AF071747 AJ011741 N85424 AL042407 AA218572 BE296748 BE083981 AL040877 AW499918 AW575045 H17813
 BE081283 AA670403 AW504327 BE094229 AA104024 A1471482 A1970337 AA737616 A1827444 AW003286 A1742333 A1344044 A1765634

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

100780 458_127

100830 4002_1

100906 4312_1

100930 16865_1

102221 3861_1

101809 32963_1

102590 15932_1

101977 29073_1

102781 20812_1

AI948838 AW253336 AW172827 AA095289 BE046383 AI734240 W16699 AI660329 AI289433 AA933778 AW469242 AA568838 AA806933
AA625873 W78031 BE206307 AA550803 AI743147 AI990075 AA948274 AA129533 AI635399 AA605313 AI624669 AW594319 AI221834 AI337434
AA307706 BE550282 AI760467 AI630636 AI221521 AW674314 AW078889 AI933732 AI686959 AI186928 AW074595 AI127486 AI079544
AI910815 H17814 AA310903 AW137854 T19279 AA026882 AA308035 AW383390 AW383389 AW383422 AW383427 AW383395 H09977
AA306247 AA352501 AW403639 F05421 AA224473 AA305321 H93904 AA089612 AW391543 AW402915 AW173382 AW402701 AW403113
R94438 N73126 H93466 AA090928 AA095051 T29025 AW951071 L47277 L47276 AI375913 BE384156 W24652 AA746288 AA568223 BE090591
H93033 N57027 AA504348 AA327653 AW959913 N53767 AA843715 AI453437 AW263710 AI076594 AA583483 AW873194 AA435535 AA236910
AI803319 AL042776 AW074313 AI887722 AI032284 AA447521 AI123885 N29334 AI354911 AW090687 AA236763 AA435535 AA236910
AA047124 AA235734 AW514610 H93467 AA952007 AI445783 AA127259 AI613495 AI686720 AI587374 AA336731 AA702453 AI859757
AA216785 AI251819 AI469227 AA806022 AI092324 N71858 AA958782 AA236919 AA809450 AA227220 AA765284 AI192007 AA766810
AA805794 AA729280 AA806238 AW768817 N71879 AI050686 AA505822 AA668974 AI688160 BE045915 AW466315 AA731314 AA649568
AA834316 AW591901 AW063876 AW294770 AI300266 AI336094 AI550380 AA721755 H09978 D20305 D29155 AW821790 BE150864 F01675
AI457474 AW466316 AA550969 AA630788
BE561958 BE561728 BE397612 BE514391 BE269037 BE514207 BE562381 BE514256 BE514403 BE514250 BE397832 BE269598 BE559865
BE396880 BE560031 BE514199 BE560037 BE560454
AC004770 W05005 AA356068 AA094281 H29358 T56781 AWB75313 L37374 BE312466 BE311755 BE207106 BE293320 BE018115 AW2239090
BE548830 AW247547 AA776082 BE397382 AA486713 T10111 T09340 AW498981 BE547280 AA356003 AW581520 AW875331 AA580720
AW875335 BE276873 BE408229 AW183148 BE255166 BE253761 AW793727 AW373141 AW581548 AA471223 AA305950 BE263976 AA526820
BE257409 AW360962 AA090655 C00312 BE312741 BE407213 AA209352 AW298199 AW248553 AW297794 AW731722 BE300586 AW731972
AW615446 BE301599 AW615520 AA485714 AW440257 AA195616 AA564630 AA618079 AW192592 AW474985 AA604580 AI627451 AA765440
AI680394 AL135548 AI683224 AI581126 AW245096 AW194154 H29274 N70363 AA629758 AA580602 AA62006 AI863841 AI097667 AI928583
AI358774 BE243487 AA620553 AA653297 AA292690 T10110 Z38906 AA908544 AA340930 AI185438 T03328 T28844 AI687010 AI684965
AI872575 BE388740 T56780 AW373138 BE258717 AA699671
AU076916 BE298110 AW239395 AW672700 NM_003875 U10860 AW651755 BE297958 C03806 AI95876 AA644165 T36030 AW392852
AA446421 AW881866 AI469428 BE548103 T96204 R94457 N78225 AI564549 AW004984 AW780423 AW675448 AW087890 AA971454 AA305698
AA879433 AA535069 AI394371 AA928053 AI378367 N59764 AI364000 AI431285 T81090 AW674657 AW674987 AA897396 AW673412 BE063175
AW674408 AI202011 R00723 AI753769 AI460161 AW079585 AW275744 AI873729 D25791 BE537646 T81139 R00722
J04129 NM_002571 AA293088 AA477016 AA404631 T28299 AA476904 AA433965 AA430486 AA495907 AI151391 AA291495 AA402723 W25551
AA706816 AI826712 AW296294 AA293479 AI276581 AW044154 AI080180 AI417985 AI274168 AI474212 AA495908 AA635664 AI092114
AI804952 AA479874 AI597661 AI420511 AA479738 AA421417 AA421247 AA436220 AL047797 M34046 N42277 AA228076 W02698 AI420297
AA434011 AI369971 AA479731 AI865541 AI418020 AA421246 AA452764 AL048051
NM_006769 U24576 AW161981 AW160473 AW160465 AW160472 AW161069 AI824831 AW162635 AI990356 AW162477 AW162571 AI520835
AW162352 AW162351 AW162752 AI952216 AI537346 AA853902 H17667 BE045346 BE559802 BE255391 AA985217 AA235051 AI129757
AW366451 T34489 D56106 D56351 AI936579 AW023219 AW889335 AW889120 AW889232 AW889175 BE093702 AW889349 AA147546
AI952998 AA912579 AI143356 AW902211 R64717 AW157236 AI815242 D45274 AW263991 AA442920 AA129965 AI035713 AI923255 AI949082
AI142826 AI684160 AI701987 AI678954 AI827349 BE463635 AW628092 AW302281 AA493203 BE348856 BE356419 AW193969 AW673561
AW592609 AI224044 H43943 AA091912 R49532 R48353 AI568409 R48256 AI198046 H27986 H43899 AI678759 AI680310 AI624220 H17052
AA155410 N56062 AI699430 AA664529 T09406 T10459 AA627508 AI379584 N83831 N86333 AW022651 AA971281 AA248036 AI039197
AI914688 AA973825 AL047305 AA129965 AI798369 AW264348 AI445879 AI658759 N59792 AI933507 AI216121 AI333174 T10972 AI375028
AI186756 AI273778 AA610487 AI797946 AA853903 AA903939 AI338587 AI274894 AW627595 AA904019
M68549 AA315280 NM_004004 AA315269 BE142653 AA461400 AW802042 BE152893 AW383155 AA490588 AW117930 AW384563 AW384544
AW384566 AW378307 AW378323 AW39085 AA257102 AW378317 AW276060 AW271245 AW378298 AW384497 AI598114 AW264544 AI018136
AW021810 AA961504 AW086214 AW771489 AW192483 AI290266 AW192488 AW384490 AW007451 AW368521 AW368463 AA461087 AI341438 AI970613
AW020066 AI783695 AI589498 AI917637 AW264471 AW384491 AI816732 AW368530 AW368521 AW368521 AW368521 AW368521 AW368521
AI040737 AI184000 AA947181 AA962716 AI280895 AW769275 AW023591 AI160977 AA055400 N71882 AA490466 AW243772 AW316635
AI076554 AW511702 N69323 H88912 AA257017 AI952506 H88913 AI912481 AA600714 BE465701 N64149 C00523 N64240 AA677120
R61573 BE005029 X8091 AI297307 BE537267 BE566138 BE566139 F11561 BE564795 BE568776 AW084005 BE566479 BE380035 BE567012
BE568634 BE566568 AA298060 BE566043 BE568813 BE568618 AA283070 BE565414 BE566738 BE568585 AA448772 AA071363 AW732642 BE564996
U62136 AF049140 BE567057 BE567297 BE567403 BE564316 BE567400 BE568854 BE566588 AA448772 AA071363 AW732642 BE564996
AA297763 AA278550 AA421083 AA298184 AA901007 AA984577 AA205916 N28759 AL031291 C15757 C15761 H02778 BE566410 AA129335
AA419499 N87741 BE379689 BE004824 BE379611 D25874 AA148454 AA323654 AW950311 AA448795 AW749423 AW773386 AA773843
AW020327 BE348580 BE504258 BE549990 BE220200 AI673334 AI202679 AA975515 D61421 AI166888 AA102843 AW246821 AI276203
AI074054 AI633824 AI962927 AI148926 N50969 AI308911 AA410994 AW373025 AA148455 H02620 AA688293 AI246318 N22220 AI917777
AI050943 AI097286 AA663794 AW368662 AW627826 AW078734 AI253060 AA749154 AA83236 AI192358 AW024576 AA448676 AA764891
BE439467 AA661534 AA258061 AI090546 AA995157 AI051011 AA584421 AI026032 AW591338 AW589563 AA776914 AW024684 AA421002
F09219 BE464500 AI383595 AA954244 AA601583 AA737304 AA195549 AA805778 AI055876 AA164942 AW013961 AI672608 AW514211 D59441
AW582574 AA160935 BE566501 BE564612 BE565353 BE568195 BE565447 BE568302 BE566097 BE565470 BE566429 AL036217 AW749424
BE567494 AA102842 AA314761 AV661237 C14211 AA651866 AW798997 AA470805
AF112213 AL050318 T24804 AW248136 BE386341 BE263177 W16577 BE250224 BE363669 BE267405 BE545577 AV651354 AV651292
AI346903 AI539128 AI189171 S63364 AW073849 AI816760 AW073309 AI422690 AA296692 AI660301 AI805446 N77735 AI340328 BE092530
AW028742 BE088442 AA657742 AA742438 AW170086 AI038920 AI432379 N36073 AI936194 AA868655 AA983612 AI077505 BE080433
AI375014 AI126547 AI348244 AI346077 AI746952 N26915 AI753574 AI093341 AI278762 BE092517 N74204 H05158 T58149 AI129303 N58366
AA524456 BE122661 AA542925 AI246120 AI735203 AA706829 AA877544 AI082289 AA926687 N92840 AW249798 AA934763 AW998363
AI128632 N25202 AI240209 AW118892 N80744 R35655 AI342321 AI340141 AW878792 AI857321 H09610 W04501 AW006650 AI126005
AA553675 AI052791 AW059835 AI041906 AA814658 AW002059 AA729483 AI609301 AA994633 AA903651 AI459183 T95072 AW088630
AA126112 AI800091 AI561215 H17502 AW475072 AI819003 AI683272 AI262701 AW793140 T81787 R95886 AI275160 AI310420 AI698929
AA159174 AI827968 F30305 F30309 AA806662 AI091923 AW878722 AA583430 AW571913 AI674584 AA292533 AI079471 AA624235 AA719050
AW793172 AA305476 AW103745 T23459 N79525 AI784438 AA534551 AW193751 AI074360 BE281214 T32229 W25066 W01205 T63086
AI795348 AI361287 AW795353 AW795349 AA594759 AI400295 D11489 AI370689 AA482356 AA485295 W40151 AA564661 AW300745
AI345938 AI374975 AI423782 AW193899 AA612604 AI183409 AA996156 AW366963 AW366977 AI284860 AA846503 AI985064 AA844576
AA737921 AA873274 BE241546 BE241540 AA484058 AW468970 AA127876 AA159120 AW001568 AW795213 AW795258 AW795330 BE250589
BE387572 AA910895 AA161217 BE250380 W31500 T95167 AI719306 AI359224
BE259778 BE281230 BE410044 T33723 AW672694 AW410439 NM_006429 AF026292 T35505 BE542333 T08940 AU076737 AW247471
BE393215 AW328640 BE542408 T32170 BE302544 T31955 BE206698 BE275738 T32570 BE386426 BE298746 BE389397 BE293991 BE315289
BE389578 R34739 R15312 BE279365 BE277758 AL036019 T33725 BE277779 BE302962 AL047294 BE276505 T09070 T33673 BE312580
AW387774 BE257175 AW574367 BE253331 BE270344 BE299831 BE273576 T32052 AI751831 BE18381 AA304899 BE252268 U46364
BE256790 BE207199 BE256209 BE251941 BE250791 BE313955 BE269806 BE543623 BE279212 BE252289 T31699 BE262220 T31669
AA315781 AA192212 N84547 BE292737 BE259631 AA232179 AI133144 T31292 AA315945 BE407301 BE251184 BE409006 AI880158 AI904003
AI904114 AW651768 AW651763 R58247 BE271897 U83843 C05298 BE261609 BE255973 AA351650 N84631 BE263637 AW452910 AA328465
AA324549 AW579525 BE252296 BE257551 AL048332 BE208630 AA359336 AW327897 AA151742 AA305816 BE076862 BE076796 BE263161
AA323785 AA676588 AA626565 AA078917 W67657 R09002 R94021 AA312032 BE276665 AA295608 AW407162 AA329374 AW877912 N27885
AA369256 AA360968 BE250476 N85427 BE265569 AI278639 AI816576 AI691037 AW325883 AI567949 AI983455 AI972732 AI811297 AI571508
AW073674 BE299039 BE467326 AI828796 AI816578 AW511604 AI921213 AW152427 AI795787 AI801618 AW168666 AI628144 AI890339
AW173690 AW511540 BE335520 AA383014 BE301164 AI866596 AW514909 AA658050 AW575243 AA074631 AI093488 AW575408 AW675443
AW615636 AW732207 AW377638 AA321784 AA641629 AA633105 AA527640 AW129145 AW815672 BE394607 AA483902 AW475032 BE378532

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

AA872808 AJ469388 AW105268 BE047301 AW591843 AW410056 AW517153 AI950495 AA746641 AI914878 AA873185 AI696911 AA548625
 AA911505 AA148762 AW574535 AI587329 BE328328 AW270348 AA158225 AW117705 AW474997 AW519193 AA614757 AW654383 AI082647
 AW590973 AI476711 AA192213 N88741 BE464552 AW072679 AA53708 AA152166 AA805924 AI581078 AIJ25768 AW173484 AI961980
 BE300766 AI199698 AI636792 AW247333 AW272861 AA078818 AA150012 AA551232 AA678821 AW873869 AW768266 AA603015 AA319210
 AA814551 AA157994 AA318885 AI582962 AW089224 AI356098 AI343694 AW072598 N21054 AI301249 AA742924 H17917 AW328584
 AW248899 AI751830 AA907816 R08898 AW087989 AI828300 AA148596 AI269577 T33426 AA213571 AI973201 AA666279 R49612 AI573183
 AW799762 AW410068 AW769566 AI962097 AI475204 D57490 AW517531 BE245270 AW470008 T33427 AW005731 AI795795 T23753 AW272981
 T15747 AA552875 T23644 AW361289 AI755559 BE207435 AA876958 T03361 AA883569 F37533 AA582321 AW082524 R42212 AA973847
 T18900 AA085202 AI599867 AI020418 AA948667 AA745670 T08939 T33724 T33722 BE621568 D57489 D25906 BE621151 F16510 C05866
 T35127 AA630427 AI933481 AA309426 AI918440 BE551854 BE618866 BE394675 BE296173 AW951687 BE383739 BE616141 BE312730
 BE535351 AW080575 BE313330 BE616664 AI354390 AA847315 BE544509 BE15212 BE297833 BE278808 BE544844 AW080178 AI890664
 BE546708 AW189943 BE274412 BE382399 BE266392 BE254949 BE280696 BE383237 BE261756 BE257721 BE312683 BE275476 BE514880
 BE545314 BE313587 BE384537 BE386691 BE264813 AW592575 AI335332 AI278641 AI795791 BE222662 AW249316 AA314361 AL036012
 AW402923 BE266845 AA075945 AA314436 BE384640 AW731769 AW957077 AA552234 AA573560 AW367038 AA313399 AI983873 BE410159
 BE263803 BE514339 BE409073 BE281296 BE543396 BE395387 BE088360 BE546946 BE546570 BE390626 AA074638 AA301821 AW845230
 AW582379 AI849222 AW029572 AA515843 AW272394 BE250234
 C14322 W74050 AI074232 AA595624 BE048955 AI148417 AI583145 AI473460 AI801688 AW573593 AI950741 AI628140 AW467921 R98105
 AI149258 AI247584 AI078378 AI139850 AA489411 W24744 R98104 AI033826 AA699589 AI033120 N55544 W88984 AW970771 AA703362
 AA099138 AA706792 AA046150 H98981 AI916674 AA953018 AI972749 AI921343 AA909044 AA094751 AI203124 AA582143 AI446654
 AW235415 R70377 AA099236 F20703 AA524436 R69484
 H04043 D60988 D60337
 AI204246 AI204250 AI194050
 H83265 T63524 AA304359 AW960551 AI672874 AI749427 AA227777 AW027055 AA971834 T49644 T54122 AI983239 AI808233 T91264 T96544
 AI350945 AI079114 R72382 T48788 AR8726 AW385418 AI095484 T49645 AA928653 AA570082 AW007545 T57178 AA516413 AA913118 T57112
 AA564433 AA774503 AA367671 T59757
 W78816 AI720806 AI633854 AI632086 AI668663 N0894 AW571809 AI383592 AI201348 W80715 N91880 AW963101 AA339011
 AB033023 BE391906 BE275965 BE277872 BE003882 AA313774 BE019159 BE298024 BE299727 BE300011 BE350277 BE384764 N87550
 BE409419 BE408652 BE408197 AL119332 AA622427 AI816265 AA610118 T07318 AD19839 AA634430 BE205794 BE04961 AI042322
 AI652711 AI917645 AA630045 AW191969 AI817882 T17271 AI803663 AJ095533 H46019 AW592438 AI624836 AI675552 D51149 AW132058
 AA639614 AI925762 AW088153 T17455 AA018640 AW751475 BE300241 AI816255 BE391981 AW408671 AA355244 AW875446 AW875744
 AW875926 AW875645 AW875647 AW38037 AL138042 AW892619 BE243018 AW995454 BE246381 BE009082 BE278921 AW967842 AA262454
 H30121
 W72062 AF088057 W76255 AI827219 AI631451 AW449295 AI354957 AI913803 T62772 AI222040 T62921 T63781
 AI678765 H12175 R14664 AI914049 AA995383 H08009 H19418 AW953728 AI358021 AA587361 AI269377 AA369905 AW957113 H27693
 AI300474 H73776 W74397 AA579604 AI130118 W72331 AI719085 AA568348 AI859045 AI814819 AI888714 BE467470 AW131258 H19419
 H27694 AI342165 AI914155 AA534872 BE018176 R60206 H11647 R45641 AI860468 BE301656 AI125453 AI498120 AA593735 AA879110
 AI016404 T35018 AA588397 AA49767 AA407365 BE501139 AA588354 AI337500 AW078532 Z41279 AI125449 AA335725 AA404338
 L42583 NM_005554 L42601 BE183076 BE541221 BE140567 L42610 V01516 J00269 AW275792 AW383052 AW380143 AI541102 BE612846
 AI541344 AW238368 BE613405 BE615705 BE615530 BE615301 AW379823 AW794706 AA194806 AA194992 AW384024 AW384000 AA641239
 AI246504 AI540333 AW238681 AA640399 AI540863 AI608860 AW862564 AW366725 AW368983 AW366870 AA596020 AW794721 AW794511
 AI591181 BE182523 AW794644 AW794620 AI935234 AI608903 AI608623 AW797060 AW084935 BE182517 BE182319 AI890082 AW238346
 AW797012 BE182522 AW794838 AI608794 AW304289 AA147193 AA595995 AW381128 AW366720 AA583718 AI828416 BE122864 AW368343
 AA431080 AW082039 AW380976 AA587144 AA443636 AW872937 AW794448 AW378382 AW085761 AW794718 AW263895 AA583587
 AA583991 AA583994 AA586886 AA586880 AW368365 AI814450 AA586991 AI282829 AW378405 AA586721 AI609242 AA431973 AA232859
 AI831095 AW263854 AW378391 AW378415 AW378381 AA036990 AW238395 AI285446 BE208219 BE049526 AA583605 AA583918 AW366711
 AI285580 AW082642 AI285712 AA582875 AW591216 AW368719 AW378408 BE122835 AA582976 BE350422 AA418328 AI541454 AI555930
 AA583700 AA150575 AW238427 AI287474 AA912658 AA584223 AW238528 C17918 AW136169 AA159847 AI923797 AI909009 BE182479
 AI915198 AW378114 AA147179 AA584239 AA150532 AW168862 AW085999 AW082480 AA659742 AW079703 AI872793 AA583981 AI824571
 BE182316 BE182507 AA233331 AI824572 AI540586 D29492 BE182931 AA036948 BE551821 D29401 AW378365 C00141 D29181 D29567
 AW103359 W95238 AI991663 AA587298 BE184608 AA099833 W95121 W95150 D29584 AI934111 D29456 D29533 AW265380 D29290
 AW238463 AA121041 D29204 AA595925 D29441 AW081840 AA587018 D29323 AA582891 BE182433 BE182437 BE182955 BE182434
 AW015534 AA314369 AA290715 BE568683 AW629494 D28364 AW995678
 AI907114 AA580734 AL041945 AA101515 AA121344 D78130 NM_003129 AA341650 T84165 AF098865 AA130976 BE089553
 T66122 AW175590 F05344 AI114790 R12900 AA194871 AA132298 D78129 AA132213 AW948930 AW948919 AA263053 AW946593 AW948840
 AA278558 R50895 N26940 N40818 AW021255 AA054851 AA663379 AW948795 AW948893 AA400356 AW948911 N85024 W78844 AI341456
 AI760182 AA288783 BE617763 BE617263 AW263690 BE049454 BE617928 AW515038 AW950584 AA601009 AI079194 AA147204 AW083163
 AA130981 AI218369 AA604784 AI808257 AI595556 AA232318 AA258065 AI471982 AA687949 AI413944 N30172 AA400196 AI769049 AI084342
 AI221380 AA948469 AJ082469 H05720 AA113270 AA158138 AA076231 AI521024 AI810962 AI133616 AA805106 AA101516 R40052 R50778
 R43280 T65036 AW131924 AA114251 AA152331 F09650 AA580614 AA558927 C75491 Z38352 AA954595 C75606 W80742
 D56165 M36981 X58965 NM_002512 BE379177 AA314836 BE256445 BE252016 AW248343 AI720933 AW085701 BE386050 BE519742
 BE277805 AA147951 AA603113 BE253293 AI246588 AI183405 AI954174 AI128891 AI829101 AI123832 AW129670 AA471268 AW170242 AW873079 AA148011 AI608620
 AA482961 AI003658 H43261 AA657978 AI735072 R83138 AA722002 AA626271 AW273977 BE464626 AA071483 AA429973 AA494342
 AA620436 AA775597 AA775601 AA826847 AI192585 AA826359 AA411159 AI193419 AI204013 AA705323 AA716255 AI784611 AI081144
 AI128227 AA829464 AI148911 AI493446 AI826084 AI189180 AI721196 AI190618 AA284987 AI128543 AA632064 AI333073 AI278470 AI131688
 AI491768 AA937581 AA630065 AA834257 AW249841 AA583742 AI309756 AA961676 AI760860 AA557818 AA954238 H43655 AI302564
 AA127545 AI909219 H20426 AI042292 AI055466 AA581836 W47002 AA422057 AA937673 F29757 AA829208 AW327462 AA372098 W02144
 AA036805 AA487365 AA961037 AI139946 AA487250 AA737118 AI952504 AI242293 AA650552 AI708401 AI633133 AA630848 AA654317 F24128
 AI434165 W46522 AW043879 AI033763 F37228 AA687809 N49087 AA876981 AA505947 AI914572 AI833284 F22253 AA026222 R50166
 AI219267 N27095 AA495512 AI784222 AI289904 AA513145 AA528547 AA418700 F36721 AI880700 AI601170 AI862851 AI708633 AA524499
 AA642220 AA495628 AI718709 W80579 AI720547 F20718 AA649943 AA588229 N40503 H46029 BE262669 BE391069 BE537538 AI510751
 AI906968 AI18611 H46099 AI472604 T60667 AA373087 W32479 AA514034 BE619183 AI134672 AA127544 H26942 BE536689 AW327461
 AA422139 AW262357 AW327348 F33510 AI630382 AW827126 F27133 AI335189 AW517599 W80471 AA895814 N89681 BE393173 AA617760
 AA584268 AA460537 AA446261 H20425 N64040 AW276801 AA316367 AA071232 BE545409 AA308292 BE274447 AA380861 AA340038
 AA341806 AA865579 AI018634 AI766314 AI919302 AA872367 AA991404 AI906961 AA888375 BE621012 AA505388 AA935192 AA290828
 R50220 H0814 H44721 AW951723 AA514795 AA418708 AW673377 AA379622 AA977995 AA708224 AA708216 AI318249 AI318233 AA441160
 AA026221 AA316774 AA466908 AI500094 AA096362 AW583742 BE536422 BE618653 R70203 AA131732 AA345048 BE562720 T28342
 NM_004415 AL031058 W77830 BE149760 AW752599 AW848723 AW376697 AW376817 AW376699 AW848371 AW376782 AW848789
 AW361413 AW849074 AW997139 AW799304 AW799309 BE077020 BE077017 BE185187 AW997196 BE156621 BE179915 BE006561 BE143155
 AW890985 BE002107 AW103521 AA857316 AW383133 BE011378 AW170253 BE185750 AW886475 BE160433 J05211 BE082676 BE082584
 BE004047 AW607238 AW377700 AW377699 BE082526 BE082505 BE082507 BE082514 AW178000 AW177933 AI905935 AW478777 AW748114
 BE148516 AW265328 AW847678 AW847688 AW365151 AW365148 AW365153 AW365156 AW365175 AW365157 AW365154 AW068840
 BE005272 AW365145 BE001925 BE182166 BE144243 BE001923 AI951766 AI434518 BE184920 BE184933 AI284090 BE184941 AW804674
 BE184924 C04715 W39488 AW995615 BE184948 BE159346 AW060653 AA099891 AA131128 AA337270 AA340777 AW384371 AA852212
 R58704 AW366566 AW364859 AA025851 AA025852 AA455100 AA719958 AW352220 AW996245 BE165351 BE073467 AA377127 AW890264
 AW609750 AW391912 AW849690 T87267 AW853812 AA852213 W74149 BE009090 AA056401 H91011 AW368529 AW390272 C18467

WO 02/086443

PCT/US02/12476

5

10

15

20

25

AW574920 N57176 AA026480 AW576767 H93284 AA026863 AW177787 AA026654 AW177786 BE092134 BE092137 BE092136 AW177784
 AI022862 BE091653 AW376811 AW848592 AA040018 BE185331 BE182164 AA368564 AW951576 T29918 AA131077 W95048 W25458
 AW205789 H90899 N29754 W32490 R20904 BE167181 BE167165 N84767 H27408 H30146 AI190590 C03378 AI554403 AI205263 AA128470
 AI392926 AF139065 AW370813 AW370827 AW798417 AW798780 AW798883 AW798569 R33557 AA149190 C03029 AW177783 AA088866
 AW370829 AA247685 BE002273 AI760816 AA39101 AW879451 AI700563 AA451923 AI340326 AI590975 T48793 AI568096 AI142882 AA039975
 AI470146 AA946936 BE057737 BE057785 W19287 AA644381 AA702424 AI417612 AI306554 AI686869 AI568892 AW190555 AI571075 AI220573
 AA056527 AI471874 AI304772 AW517828 AI915596 AI627383 AI270345 AW021347 AW166807 AW105614 AI346078 AA552300 W95070
 AI494069 AI911702 AA149191 AA026864 AI830049 AI887258 AW780435 AI910434 AI819984 AI858282 AI078449 AI025932 AI860584 AI635878
 AA026047 AA703232 D12062 AW192085 AA658154 AW514597 AW591892 T87181 AA782066 AW243815 AW150038 AW268383 AW004633
 AI927207 AA782109 AW473233 AI804485 AW169216 AI572669 AA602182 AW015480 AW771855 AI270027 AA961816 AA283207 AI076962
 AI498487 AI348053 AI783914 H44405 AW799118 AA128330 AA515500 AA918281 W02156 AI905927 AA022701 W38382 R20795 T77861
 AW860878
 100528 45979_1 BE386801 AU077299 AA143755 BE302747 AA853375 U30162 BE274163 BE277479 BE408180 BE274874 C15000 AA047476 N27099 AI359165
 AI638794 AI151283 AI863925 AW444977 AI207392 AA931263 AA443112 R40138 AW068538 AA351008 AA676972 R62503 AA916492 AW001865
 H42334 H38280 AA121497 AA114137 AI750938 M17783 AA383785 BE274462 AI753182 C05975 AA347404 AW069298 AI754351 AI754044
 AA188808 AA186879 AA565243 AL040655 AA456177 AI750722 AA045756 AA213580 C16936 AW578747 AW753731 H41632 N44761 R58560
 R61260 AA039902 N59721 AW992543 R68380 AA149686 T29017 H03739 BE383822 BE387105 BE408251 BE410425 H41560 AA247591
 BE389677 AI752233 AI566195 AA868004 AI424523 AW753720 AA852159 BE386803
 100559 2260_1 NM_000094 L02870 D13694 S51236 M96984 AW946290 M65158 AI285422 D29523 AL119886 AW630655 L06862 AI884355 AW168737 T29085
 AW797005 AW801340 AI355504 AW079048 AW801337 AI690455 AI972063 AW268565 W68583 AA587326 AA883498 AI033523 AW510356
 AW591998 H93463 AL043852 AI150055 AI566239 AI624803 AA844717 H40670 AA922334 AI864424 AW615094 AW451233 AI302203 F31221
 AI872170 W68589 AA904478 AI917631 AW014208 AW450759 AA847625 AI284033 AA848176 AA598507
 100576 9986_1 X00356 NM_001741 M26095 X03662 M12667 X02330 X02330 AA716058 AW296074 X04861 AI695720 AA719597
 124357 genbank_N22401 N22401
 101624 entrez_M55998 M55998
 101625 entrez_M57293 M57293
 135158 57963_1 AL037551 AI804716 AW439811 AI559470 AA075299 AI738572 AI270388 AI816783 AW263026 AI633951 AI655285 AI990572 AI950425
 AW241533 AA916883 AA576693 AA160156 AA613783 AW078884 AI888282 AI275241 AI133467 AA164921

Tables 2A-8C were previously filed on November 9, 2001 in USSN 60/339,245 (18501-004100US)

Table 2A shows 504 genes down-regulated in lung tumors relative to normal lung and chronically diseased lung. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number										
	ExAccn:	Exemplar Accession number, Genbank accession number										
10	UnigeneID:	Unigene number										
	Unigene Title:	Unigene gene title										
	R1:	90th percentile of AI for normal lung samples divided by the 80th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples.										
	R2:	median of AI for normal lung samples divided by 90th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples.										
15	R3:	median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.										
	R4:	average of AI for normal lung samples divided by average AI for squamous cell carcinoma and adenocarcinoma lung tumors.										
	R5:	median of AI for normal lung samples divided by the 90th percentile of AI for adenocarcinomas.										
20	R6:	median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for adenocarcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.										
	R7:	average of AI for normal lung samples divided by the 90th percentile of AI for squamous cell carcinomas.										
	R8:	median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for squamous cell carcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.										
25	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2	R3	R4	R5	R6	R7	R8
	100095	Z97171	Hs.78454	myocilin; trabecular meshwork inducible	40.20							
	100115	NM_002084	Hs.336920	glutathione peroxidase 3 (plasma)								3.46
30	100138	U83508	Hs.2463	angiopoietin 1			2.30					
	100299	D49493	Hs.2171	growth differentiation factor 10		11.00						
	100306	U85749	Hs.80598	transcription elongation factor A (SII);						3.06		
	100447	NM_014767	Hs.74583	KIAA0275 gene product								3.16
	100458	S74019	Hs.247979	Vpre-B	42.40							
	100862	AA005247	Hs.285754	Hepatocyte Growth Factor Receptor						4.13		
35	100959	AA359129	Hs.118127	actin; alpha; cardiac muscle				125.60				
	101032	BE205854	Hs.46039	phosphoglycerate mutase 2 (muscle)	36.40							
	101081	AF047347	Hs.4880	amyloid beta (A4) precursor protein-bind				34.60				
	101088	X70697	Hs.553	solute carrier family 6 (neurotransmitter)				193.20				
	101125	AJ250562	Hs.82749	transmembrane 4 superfamily member 2						3.10		
40	101180	U11874	Hs.846	interleukin 8 receptor; beta				54.86				
	101308	L41390		"Homo sapiens core 2 beta-1,6-N-acetylgl	33.20							
	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do				36.40				
	101345	NM_005795	Hs.152175	Calcitonin receptor-like			2.29					
	101346	AJ738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N				70.55				
45	101397	M26380	Hs.180878	lipoprotein lipase							3.81	3.54
	101414	NM_000066	Hs.38069	complement component 8; beta polypeptide								
	101435	NM_001100	Hs.1288	actin; alpha 1; skeletal muscle				34.60				
	101507	X16896	Hs.82112	interleukin 1 receptor; type I				37.60				
	101530	M29874	Hs.1360	cytochrome P450; subfamily IIB (phenobar								4.25
50	101537	AI469059	Hs.184915	zinc finger protein; Y-linked			2.54					
	101542	NM_000102	Hs.1363	cytochrome P450; subfamily XVII (steroid		5.50						
	101545	BE246154	Hs.154210	EDG1; endothelial differentiation, sphin	39.40							
	101554	BE207611	Hs.123078	thyroid stimulating hormone receptor		13.00						
	101560	AW958272	Hs.83733	intercellular adhesion molecule 2, exon								3.38
55	101574	M34182	Hs.158029	protein kinase; cAMP-dependent; catalyti						4.37		
	101605	M37984	Hs.118845	tropotin C; slow								3.80
	101621	BE391804	Hs.62661	guanylate binding protein 1; interferon-	30.20							
	101680	AA299330	Hs.1042	Sjogren syndrome antigen A1 (SZK); ribon							2.75	
	101829	AW452398	Hs.129763	solute carrier family 8 (sodium/calcium						3.37		
60	101842	M93221	Hs.75182	mannose receptor; C type 1				38.20				
	101961	AW004056	Hs.168357	"Hs-TBX2=T-box gene (T-box region) [huma			2.32					6.85
	101994	T92248	Hs.2240	ulceroglobulin								
	102020	AU077315	Hs.154970	transcription factor CP2			2.45					6.75
65	102091	BE280901	Hs.83155	aldehyde dehydrogenase 7								
	102112	AW025430	Hs.155591	forkhead box F1	54.60							3.98
	102190	AA723157	Hs.73769	folate receptor 1 (adult)								3.62
	102202	NM_000507	Hs.574	fructose-bisphosphatase 1								
	102241	NM_007351	Hs.268107	Multimerin			2.32					
70	102310	U33839		Accession not listed in Genbank		7.00						
	102397	U41898		"Human sodium cotransporter RKST1 mRNA,	29.40							
	102571	U60115	Hs.239069	"Homo sapiens skeletal muscle LIM-protei								3.75
	102620	AA976427	Hs.121513	Human clone W2-6 mRNA from chromosome X						3.07		
	102636	U67092		"Human ataxia-telangiectasia locus prote			2.40					
	102667	U70867	Hs.83974	solute carrier family 21 (prostaglandin			3.15					
75	102675	U72512	Hs.7771	"Human B-cell receptor associated protei						3.56		
	102698	M18667	Hs.1867	progastrin (pepsinogen C)								4.51
	102727	U79251	Hs.99902	opioid-binding protein/cell adhesion mol					12.00			
	102852	V00571	Hs.75294	corticotropin releasing hormone	37.40							
	103026	X54162	Hs.79386	thyroid and eye muscle autoantigen D1 (6					13.00			
80	103028	X54380	Hs.74094	pregnancy-zone protein	28.80							
	103098	M86361		Human mRNA for T cell receptor; clone IG					10.00			
	103117	X63578	Hs.295449	parvalbumin		6.00						
	103241	X76223		H.sapiens MAL gene exon 4			2.47					
	103280	U84722	Hs.76206	Cadherin 5, VE-cadherin (vascular epithe			2.69					
85	103360	Y16791	Hs.73082	keratin; hair; acidic; 5							2.16	

WO 02/086443

PCT/US02/12476

	103496	Y09267	Hs.132821	flavin containing monooxygenase 2				3.27	5.97
	103508	Y10141		*Hsapiens DAT1 gene, partial, VNTR					
	103561	NM_001843	Hs.143434	contactin 1	2.40				
5	103569	NM_005512	Hs.151641	glycoprotein A repetitions predominant	2.99			4.18	
	103575	Z26256		*Hsapiens isoform 1 gene for L-type cal				3.44	
	103627	Z48513		Hsapiens XG mRNA (clone PEP6)					2.25
	103767	BE244667	Hs.296155	CGI-100 protein		46.55			
	103850	AA187101	Hs.213194	Hypothetical protein MGC10895; sim to SR				3.05	
10	104078	AA402801	Hs.303276	ESTs				3.54	
	104326	AW732858	Hs.143067	ESTs				3.16	
	104352	BE219898	Hs.173135	dual-specificity tyrosine-(Y)-phosphoryl					
	104398	AI423930	Hs.36790	ESTs; Weakly similar to putative p150 [H	64.80			3.38	
	104473	AI904823	Hs.31297	ESTs					
	104493	AW950427	Hs.79059	ESTs; Moderately similar to TGF-BETA REC		2.47			
15	104495	AW975687	Hs.292979	ESTs	28.60			3.42	
	104595	AI799603	Hs.271568	ESTs					
	104597	AI354504	Hs.93957	ESTs; Weakly similar to Slt1-1 protein [6.00			
	104659	AW959769	Hs.105201	ESTs	34.00				
	104686	AA010539	Hs.18912	ESTs		11.00			
20	104691	U29690	Hs.37744	ESTs; Beta-1-adrenergic receptor	58.80				
	104764	AI039243	Hs.278585	ESTs		60.40			
	104776	AA026349		ESTs	34.20				
	104825	AA035613	Hs.141883	ESTs		3.03			
	104865	T79340	Hs.22575	Homo sapiens cDNA: FLJ21042 fis, clone C	41.20				3.27
25	104942	NM_016348	Hs.10235	ESTs					
	104989	R65998	Hs.285243	ESTs		40.00			3.20
	105062	AW954355	Hs.36529	ESTs	34.20				
	105101	H63202	Hs.38163	ESTs					4.17
	105173	U54617	Hs.8364	ESTs					
30	105194	R06780	Hs.19800	ESTs		16.00			
	105226	R58959	Hs.26608	ESTs			2.34		
	105256	AA430650	Hs.16529	transmembrane 4 superfamily member (tet			2.72		
	105394	BE245812	Hs.8941	ESTs			2.61		
	105647	Y09306	Hs.30148	homodomain-interacting protein kinase 3	33.60				3.59
35	105789	AF106941	Hs.18142	arrestin; beta 2				4.46	
	105817	AA397825		synaptopodin					
	105847	AW964490	Hs.32241	ESTs		35.40			
	105894	AI904740	Hs.25691	calcitonin receptor-like receptor activi		3.43			
	105999	BE268786	Hs.21543	ESTs	7.00				
40	106075	AA045290	Hs.25930	ESTs		42.60			
	106178	AL049935	Hs.301763	KIA0554 protein	34.80				
	106381	AB040916	Hs.24106	ESTs		12.00		3.69	
	106467	AA450040	Hs.154162	ADP-ribosylation factor-like 2			96.40		
	106536	AA329648	Hs.23804	ESTs			47.20		
45	106569	R20909	Hs.300741	sorcin			220.40		
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr					
	106842	AF124251	Hs.26054	novel SH2-containing protein 3		2.55			
	106844	AA485055	Hs.158213	sperm associated antigen 6	39.20				
	106870	AI983730	Hs.26530	serum deprivation response (phosphatidyl		2.28			
50	106943	AW888222	Hs.9973	ESTs				4.28	
	106954	AF128847	Hs.204038	ESTs				4.32	
	107106	AA862496	Hs.28482	ESTs			10.45		
	107163	AF233588	Hs.27018	ESTs		2.57			
	107201	D20378	Hs.30731	EST				3.84	
55	107238	D59362	Hs.330777	EST		8.00			
	107376	U90545	Hs.327179	solute carrier family 17 (sodium phospho		10.67			
	107530	Y13622	Hs.85087	latent transforming growth factor beta b		2.32			
	107688	AW082221	Hs.60536	ESTs			34.60		
	107706	AA015579	Hs.29276	ESTs	28.40				
60	107723	AA015967		EST				3.29	
	107727	AA149707	Hs.173091	DKFZP434K151 protein			80.80		
	107750	AA017291	Hs.60781	ESTs			51.40		
	107751	AA017301	Hs.235390	ESTs				3.14	
	107873	AK000520	Hs.143811	ESTs		9.00			
65	107899	BE019261	Hs.83869	ESTs; Weakly similar to III ALU SUBFAM I				3.65	
	107994	AA036811	Hs.48469	ESTs			44.60		
	107997	AL049176	Hs.82223	Human DNA sequence from clone 141H5 on c			32.00		
	108041	AW204712	Hs.61957	ESTs			30.80		
	108048	AI797341	Hs.165195	ESTs				4.75	
70	108338	AA070773		*zm53g11.s1 Stratagene fibroblast (#9372		2.33		2.92	
	108434	AA078899		*zm94b1.s1 Stratagene colon HT29 (#93722					
	108447	AA079126		*zm92a11.s1 Stratagene ovarian cancer (#				3.05	
	108480	AL133092	Hs.68055	ESTs			34.00		
	108499	AA083103		*zn1b12.s1 Stratagene hNT neuron (#93723					3.36
75	108535	R13949	Hs.226440	Homo sapiens clone 24881 mRNA sequence			19.00		
	108550	AA084867		*zn11f6.s1 Stratagene hNT neuron (#93723			12.00		
	108604	AA934589	Hs.49696	ESTs		2.33			
	108625	AW972330	Hs.283022	ESTs				3.42	
	108629	AA102425		*zn24c6.s1 Stratagene neuroepithelium NT					
80	108655	AA099960		*zm65c6.s1 Stratagene fibroblast (#93721		7.00			
	108756	AA127221	Hs.117037	Homo sapiens mRNA; cDNA DKFZp564N1164 (f		6.05			
	108864	AI733852	Hs.199957	ESTs	28.80				
	108895	AL138272	Hs.62713	ESTs	32.80				
	108921	AI568801	Hs.71721	ESTs			57.80		
85	108967	AA142989	Hs.71730	ESTs	28.80				

PCT/US02/12476

97

WO 02/086443

PCT/US02/12476

	114776	AA151719	Hs.95834	ESTs	34.40				
	115009	AA251561	Hs.48689	ESTs	30.20				
	115272	AW015347		ESTs; Weakly similar to hypothetical L1	32.60				
5	115279	AW964897	Hs.290825	ESTs		6.00			
	115302	AL109719	Hs.47578	ESTs			12.00		
	115365	AW976252	Hs.268391	ESTs				3.32	
	115559	AL079707	Hs.207443	ESTs			48.00		
	115566	AI142336	Hs.43977	ESTs			56.20		
10	115883	AF255910	Hs.54650	ESTs; Weakly similar to (define not ava	31.40				
	115744	AA418538	Hs.43945	ESTs; Highly similar to AJ1178H5.3 [Hsa			33.60		
	115819	AA486620	Hs.41135	Endomucin 2			74.40		
	115949	AI478427	Hs.43125	ESTs		3.18			
	115965	AA001732	Hs.173233	ESTs			388.80		
15	116035	AA621405	Hs.184664	ESTs			33.20		
	116049	AA454033	Hs.41644	ESTs			45.80		
	116081	AI190071	Hs.55278	ESTs				3.57	
	116082	AB029496	Hs.59729	ESTs		3.06			
	116213	AA292105	Hs.326740	leucine rich repeat (in FLII) interactin	50.60				
20	116228	AI767947	Hs.50841	ESTs; Weakly similar to tuftelin [M.musc		3.85			
	116250	N76712	Hs.44829	ESTs		6.00			
	116419	AI613480	Hs.47152	ESTs; Weakly similar to testicular tek1			30.00		
	116617	D80761	Hs.45220	EST		2.27			
	116784	AB007979	Hs.301281	tenascin R (restrictin; janusin)	47.20				
25	116835	N39230	Hs.38218	ESTs			41.20		
	116970	AB023179	Hs.9059	KIAA0962 protein			91.00	11.00	
	117023	AW070211	Hs.102415	ESTs					
	117027	AW085208	Hs.130093	ESTs	49.40		32.60		
	117036	H88908	Hs.41192	EST					
30	117110	AA160079	Hs.172932	ESTs		8.67	30.60		
	117209	W03011	Hs.306881	ESTs				9.29	
	117325	N23599	Hs.43396	ESTs					3.19
	117454	N29569	Hs.44055	ESTs					
	117475	N30205	Hs.93740	ESTs	44.00				
35	117543	BE219453	Hs.42722	ESTs		16.00			
	117567	AW444761	Hs.44565	ESTs			12.00		
	117570	N48649	Hs.44583	ESTs			11.00		
	117600	N34963	Hs.44676	EST					3.74
	117730	N45513	Hs.46808	ESTs		6.00			
	117791	N48325	Hs.93956	EST		9.00			
40	117929	N51075	Hs.47191	ESTs			29.20		
	117990	AA446167	Hs.47385	ESTs		8.00			
	118224	N62275	Hs.48503	EST	31.40				
	118244	N62516	Hs.48556	ESTs	32.80				
45	118357	AL109667	Hs.124154	Homo sapiens mRNA full length Insert cDN		2.40			
	118446	N66361	Hs.269121	ESTs		2.28			
	118447	N66399	Hs.49193	EST	30.80				
	118530	N67900	Hs.118446	ESTs				3.10	
	118549	N68163	Hs.322954	EST				3.41	
	118823	W03754	Hs.50813	ESTs; Weakly similar to long chain fatty		3.94			
50	118862	W17065	Hs.54522	ESTs			33.00		
	118935	AI979247	Hs.247043	KIAA0525 protein				11.43	
	118944	AI734233	Hs.226142	ESTs; Weakly similar to !!!! ALU SUBFAM1		14.00			
	118995	N94591	Hs.323056	ESTs			52.60		
55	119073	BE245360	Hs.279477	ERG-2/ERG-1; V-ets avian erythroblastosi					
	119268	T16335	Hs.65325	EST	31.40				3.50
	119514	W37937		Accession not listed in Genbank					
	119824	W74536	Hs.184	advanced glycosylation end product-speci		2.75			3.21
	119831	AL117654	Hs.58419	DKFZP586L2024 protein			33.80		
60	119861	W78816	Hs.49943	ESTs; Moderately similar to !!!! ALU SUB			30.03		
	119889	W84346	Hs.58671	ESTs					
	119921	W86192	Hs.58815	ESTs	29.00				3.80
	120082	H80286	Hs.40111	ESTs					
	120094	AA811339	Hs.124049	ESTs		6.00			
65	120132	W57554	Hs.125019	Human lymphoid nuclear protein (LAF-4)			36.60		
	120378	AA223249	Hs.285728	ESTs		12.00			
	120404	AB023230	Hs.96427	KIAA1013 protein	39.40				
	120504	AA256837		ESTs			8.00		
	120512	N55761	Hs.194718	ESTs	33.00				
70	120567	AA287740	Hs.78335	microtubule-associated protein; RP/EB fa				4.18	
	120777	AA287702	Hs.10031	KIAA0955 protein			46.60		
	121082	AA398722		ESTs			39.00		
	121191	AA400205	Hs.104447	ESTs	41.60				
	121248	AA400914	Hs.97827	EST				5.08	
	121363	AI287280	Hs.97933	ESTs			12.00		
75	121366	AI743515		ESTs			20.00		
	121483	AI660332	Hs.25274	ESTs; Moderately similar to putative sev				3.32	
	121518	AA412155		ESTs			30.20		
	121545	AA412442	Hs.98132	ESTs		2.29			
80	121622	AA416931	Hs.126065	ESTs		9.00			
	121665	AA416556	Hs.98234	ESTs			34.80		
	121709	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	34.80				
	121730	AI140583	Hs.98328	ESTs	38.80				
	121740	AA421138	Hs.98334	EST		7.00			
	121772	AI590770	Hs.110347	Homo sapiens mRNA for alpha Integrin bin	36.20				
85	121821	AL040235	Hs.3346	ESTs					3.61

WO 02/086443

PCT/US02/12476

	121835	AB033030	Hs.300670	ESTs		2.34			
	121841	AA427794	Hs.104864	ESTs		2.61			
	121885	AA934883	Hs.98467	ESTs				2.25	
	121888	AA426429	Hs.98463	ESTs				2.92	
5	121938	AA428659	Hs.98610	ESTs			46.80		
	121950	AA429515		EST			31.40		
	122030	AA431310	Hs.98724	ESTs	34.40				
	122054	AA431725	Hs.98746	EST				3.58	
10	122211	AA300900	Hs.98849	ESTs; Moderately similar to bithoraxoid-	49.40				
	122233	AA436455	Hs.98872	EST	29.80				
	122247	AA436576	Hs.98890	EST			39.80		
	122253	AA436703	Hs.104936	ESTs; Weakly similar to hypothetical pro		9.00			
	122266	AA436840	Hs.98907	EST				3.60	
	122285	AA436981	Hs.121602	EST				3.14	
15	122409	AA446830	Hs.99081	ESTs	30.80				
	122485	AA524547	Hs.160318	phospholemman		2.65			
	122697	AA420683	Hs.98321	Homo sapiens cDNA FLJ14103 fis, clone MA		15.00			
	122772	AW117452	Hs.99489	ESTs		6.67			
20	122831	AI857570	Hs.5120	ESTs				3.37	
	122913	AI638774	Hs.105328	ESTs			32.20		
	123049	BE047680	Hs.211869	ESTs			41.80		
	123076	AI345569	Hs.190046	ESTs	35.80				
	123136	AW451999	Hs.194024	ESTs				2.58	
25	123309	N52937	Hs.102679	ESTs			19.00		
	123455	AA353113	Hs.112497	ESTs			82.80		
	123691	AA609579	Hs.112724	ESTs				3.95	
	123756	AA609971	Hs.112795	EST	35.40				
	123802	AA620448		Homo sapiens clone 24760 mRNA sequence	58.00				
30	123837	AI807243	Hs.112893	ESTs			32.40		
	123844	AA938905	Hs.120017	olfactory receptor, family 7; subfamily		2.63			
	123936	NM_004673	Hs.241519	ESTs	29.00				
	123987	C21171	Hs.95497	ESTs; Weakly similar to GLUCOSE TRANSPOR			70.60		
	124013	AI521936	Hs.107149	ESTs; Weakly similar to PTB-ASSOCIATED S	28.40				
35	124160	R40290	Hs.124685	ESTs			13.00		
	124205	H77570	Hs.108135	ESTs				4.74	
	124226	AA618527	Hs.190266	ESTs		2.35			
	124246	H67680	Hs.270962	ESTs			29.40		
	124348	AI796320	Hs.10299	ESTs		17.00			
40	124358	AW070211	Hs.102415	*yw35g11.s1 Morton Fetal Cochlea Homo sa		3.07			
	124409	AI814166	Hs.107197	ESTs				3.14	
	124442	AW663632	Hs.285625	TATA box binding protein (TBP)-associate		2.48			
	124468	N51413	Hs.109284	ESTs			30.80		
	124479	AB011130	Hs.127436	calcium channel; voltage-dependent; alph				6.03	
45	124519	AI670056	Hs.137274	ESTs; Weakly similar to SPLICEOSOME ASSO		2.50			
	124711	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	59.20				
	124866	AI768289	Hs.304389	ESTs		8.00			
	124874	BE550182	Hs.127826	ESTs			37.60		
	125097	AW576389	Hs.335774	ESTs			10.00		
50	125179	AW205468	Hs.103118	ESTs				3.12	
	125200	AW836591	Hs.103156	ESTs				2.79	
	125299	T32982	Hs.102720	ESTs			34.20		
	125400	AL110151	Hs.128797	DKFZP586D0824 protein	29.00				
	125810	H00083		aryl hydrocarbon receptor-interacting pr	32.20				
55	126176	BE242256	Hs.2441	KIAA0022 gene product		12.00			
	126303	D78841		HUM525A05B Human placenta polyA+ (TFuji			33.60		
	126403	AW629054	Hs.125976	ESTs; Weakly similar to metalloprotease/	35.80				
	126507	AL040137	Hs.23964	ESTs; Weakly similar to HC1 ORF [M.muscu			29.80		
	126773	AA648284	Hs.187584	ESTs					
60	127307	AW962712	Hs.126712	ESTs; Weakly similar to pL2 hypotheticala	39.60				
	127462	AA760776	Hs.293977	aa59b04.s1 NCI_CGAP_GCB1 Homo sapiens c	28.80				
	127486	AW002846	Hs.105468	ESTs		9.00			
	127572	AA594027	Hs.191788	ESTs		2.36			
	127609	X80031	Hs.530	ESTs			29.40		
65	127832	AW976035	Hs.292396	ESTs			37.20		
	127898	AA774725	Hs.128970	ESTs				4.42	
	128073	AW340720	Hs.125983	ESTs			38.40		
	128101	AA905730	Hs.128254	ESTs		7.33			
	128149	NM_012214	Hs.177576	mannosyl (alpha-1,3)-glycoprotein beta-				2.58	
70	128212	W27411	Hs.336920	glutathione peroxidase 3 (plasma)		3.09			
	128333	W68800	Hs.12126	ESTs; Weakly similar to LR8 [H.sapiens]			34.40		
	128364	N76462	Hs.269152	ESTs; Weakly similar to ZINC FINGER PROT		10.00			
	128426	AI265784	Hs.145197	ESTs				4.31	
	128598	AA305407	Hs.102308	potassium inwardly-rectifying channel; s	31.20				
75	128634	AA464918		ESTs; Moderately similar to !!!! ALU SUB			41.60		
	128687	AW271273	Hs.23767	ESTs			87.00		
	128726	AI311238	Hs.104476	ESTs				4.02	
	128773	NM_004131	Hs.1051	granzyme B (granzyme 2; cytotoxic T-lymp			9.00		
	128833	W26667	Hs.184581	ESTs				3.76	
80	128870	H39537	Hs.75309	eukaryotic translation elongation factor		2.66			
	128878	R25513	Hs.10583	ESTs				3.10	
	128885	AF134803	Hs.180141	cofilin 2 (muscle)			11.00		
	128998	W04245	Hs.107761	ESTs; Weakly similar to PUTATIVE RHO/RAC				3.21	
	129000	AA744902	Hs.107767	ESTs; Moderately similar to CaM-KII inh1				3.68	
85	129038	AW156903	Hs.108124	ribosomal protein L41				3.17	
	129098	AW580945	Hs.330466	ESTs	34.60				

	129210	AL039940	Hs.202949	KIAA1102 protein					4.09
	129240	AA361258	Hs.237868	interleukin 7 receptor	2.29				
	129262	BE222198	Hs.109843	ESTs			3.30		
5	129301	AF182277	Hs.330780	Human cytochrome P450-11B (h11B3) mRNA;					4.05
	129331	AW167668	Hs.279772	ESTs; Highly similar to CGI-38 protein [4.09
	129381	AW245805	Hs.110903	claudin 5 (transmembrane protein deleted	2.93				
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1		160.80			
	129595	U09550	Hs.1154	oviductal glycoprotein 1; 120kD			10.00		
10	129613	AW978517	Hs.172847	ESTs; Weakly similar to collagen alpha 1			3.40		
	129782	AW016932	Hs.104105	EST	9.00				
	129950	F07783	Hs.1369	decay accelerating factor for complement					
	129958	R27496	Hs.1378	annexin A3		87.80			
	129959	AL036554	Hs.274463	defensin; alpha 1; myeloid-related seque	2.72	44.60			
15	130160	AA305688	Hs.267695	UDP-Gal:betaGlcNAc beta 1,3-galactosyltr		42.20			
	130259	NM_000328	Hs.153614	retinitis pigmentosa GTPase regulator	2.54				
	130273	AW972422	Hs.153863	MAD (mothers against decapentaplegic; Dr		51.60			
	130312	AF056195	Hs.15430	DKFZP586G1219 protein			3.16		4.11
	130436	NM_001928	Hs.155597	D component of complement (adipsin)			4.77		
20	130523	AA999702	Hs.214507	ESTs					
	130799	AB028945	Hs.12696	ESTs	6.00		3.54		
	130885	NM_005883	Hs.20912	adenomatous polyposis coli like					
	131002	AL050295	Hs.22039	KIAA0758 protein					3.50
	131012	AL039940	Hs.202949	KIAA1102 protein	20.00				
25	131031	NM_001650	Hs.288650	aquaporin 4	41.20				
	131061	NS4328	Hs.268744	ESTs; Moderately similar to KIAA0273 [H.		31.40			
	131066	AW169287	Hs.22588	ESTs		29.60			
	131082	AJ091121	Hs.246218	ESTs; Weakly similar to zinc finger prot			9.00		
	131087	AF147709	Hs.22824	ESTs; Weakly similar to p160 myb-binding					3.86
30	131161	AF033382	Hs.23735	potassium voltage-gated channel; subfam			3.14		
	131179	AA171368	Hs.184482	DKFZP586D0624 protein			3.80		
	131182	AI824144	Hs.23912	ESTs					3.67
	131205	NM_003102	Hs.2420	superoxide dismutase 3; extracellular	2.98				
	131277	AA131466	Hs.23767	ESTs	3.15				
35	131281	AA251716	Hs.25227	ESTs		32.20			
	131282	X03350	Hs.4	alcohol dehydrogenase 3 (class I); gamma					3.44
	131285	AI567943	Hs.25274	ESTs; Moderately similar to putative sev			6.40		
	131355	R52804	Hs.25956	DKFZP564D206 protein	8.00				
	131391	AW085781	Hs.26270	ESTs	10.00				
40	131461	AA992841	Hs.27263	butyrate response factor 2 (EGF-response	28.80				
	131487	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f					4.03
	131517	AB037789	Hs.263395	ESTs; Highly similar to semaphorin Via [39.00				
	131545	AL137432	Hs.28564	ESTs			11.00		
	131563	AK000383	Hs.323092	ESTs; Weakly similar to dual specificity			10.00		
45	131647	AA359615	Hs.30089	ESTs	2.47				
	131675	H15205	Hs.30509	ESTs			3.06		
	131676	AI126821	Hs.30514	ESTs					
	131708	S60415	Hs.30941	calcium channel; voltage-dependent; beta	45.80				
	131717	X94630	Hs.3107	CD97 antigen	2.28				3.78
50	131756	AA443966	Hs.31585	ESTs		40.60			
	131762	AA744902	Hs.107767	ESTs; Moderately similar to CaM-KII inh					3.67
	131821	AA017247	Hs.164577	ESTs	2.87				
	131839	AB014533	Hs.33010	KIAA0633 protein					3.48
	131861	AL095858	Hs.184245	KIAA0929 protein Msx2 interacting nuclea	54.00				
55	132015	AI418006	Hs.3731	ESTs		49.20			
	132070	BE622641	Hs.38489	ESTs		34.80			
	132242	AA332697	Hs.42721	ESTs	2.68				
	132334	AW080704	Hs.45033	lacrima proline rich protein	4.66				
	132476	AL119844	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regl	34.20				
60	132490	NM_001290	Hs.4980	UIM binding domain 2					
	132533	AJ922988	Hs.172510	ESTs	13.00				
	132598	X80031	Hs.530	collagen; type IV; alpha 3 (Goodpasture		30.60			
	132619	H28855	Hs.53447	ESTs; Moderately similar to kinesin lig					4.02
	132652	N41739	Hs.61260	ESTs					3.18
	132726	N52298	Hs.55608	ESTs; Weakly similar to cDNA EST yk484g1			11.43		
65	133028	R51604	Hs.300842	ESTs	2.37				
	133071	BE334932	Hs.64313	ESTs	2.27				
	133120	NM_003278	Hs.65424	tetranectin (plasminogen-binding protein	2.63				
	133129	AA428580	Hs.65551	ESTs					5.49
	133147	AA026533	Hs.66	interleukin 1 receptor-like 1	6.20				
70	133151	NM_014051	Hs.94896	ESTs			3.69		
	133213	AA903424	Hs.6786	ESTs		31.40			
	133276	AW978439	Hs.69504	ESTs			9.00		
	133377	AJ131245	Hs.7239	SEC24 (S. cerevisiae) related gene famil	41.20				
75	133407	AF017987	Hs.7306	secreted frizzled-related protein 1	50.20				
	133535	AL134030	Hs.284180	protocadherin 2 (cadherin-like 2)			3.72		
	133537	U41518	Hs.74602	aquaporin 1 (channel-forming integral pr					3.35
	133656	BE149455	Hs.75415	Accession not listed in Genbank	2.65				
	133689	NM_001872	Hs.75572	carboxypeptidase B2 (plasma)		90.80			
80	133779	T58486	Hs.222566	ESTs			3.05		
	133978	AF035718	Hs.78061	transcription factor 21	2.92				
	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec					3.45
	134000	AW175787	Hs.334841	selenium binding protein 1					4.05
	134111	AI372588	Hs.8022	TU3A protein	4.49				
	134185	AA285136	Hs.301914	Homo sapiens mRNA; cDNA DKFZp586K1220 (f					3.27
85	134204	AI873257	Hs.7994	ESTs; Weakly similar to CGI-69 protein [40.80			

WO 02/086443

PCT/US02/12476

	134641	AI092634	Hs.156114	protein tyrosine phosphatase; non-recept					
	134677	AA251363	Hs.177711	ESTs			32.20		3.76
	134745	NM_000685	Hs.89472	angiotensin receptor 1B	15.00				
5	134749	T28499	Hs.89485	carbonic anhydrase IV		3.05			
	134786	T29618	Hs.89640	angiotensin 1 receptor; TEK tyrosine ki			57.80		
	134825	U33749	Hs.197764	thyroid transcription factor 1				3.73	
	134978	AJ829008	Hs.333383	foolin (collagen/fibrinogen domain-cont		2.52			
	135010	N50465	Hs.92927	ESTs			31.60		
	135053	AW796190	Hs.93678	ESTs				3.21	
10	135081	AF069517	Hs.173993	RNA binding motif protein 6	28.80				
	135091	AA493650	Hs.94357	ESTs				4.24	
	135135	AA775910	Hs.95011	syntrophin; beta 1 (dystrophin-associat		8.00			
	135203	C15737	Hs.269386	ESTs				4.31	
	135236	AI636208	Hs.96901	ESTs	43.00				
15	135266	R41179	Hs.97393	Human mRNA for KIAA0328 gene; partial cd				6.42	
	135346	NM_000928	Hs.992	phospholipase A2; group IB (pancreas)		3.82			
	135378	AW961818	Hs.24379	potassium voltage-gated channel; shaker-		4.15			
	135387	NM_001972	Hs.99863	elastase 2; neutrophil	37.20				
	135388	W27965	Hs.99865	EST	38.80				
20	135402	L12398	Hs.99922	dopamine receptor D4				4.21	

25 TABLE 2B shows the accession numbers for those primekeys lacking unigenelD's for Table 2A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

30 Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
35	108447	43452_-7 AA079126
	108550	120073_1 AA084867 AA084996
	108655	127522_1 AA099960 AA113013
	102397	44371_-1 U41898
	126303	1525933_1 D78841 D78880
40	125810	1554054_1 H00083 R81062
	103627	2615_2 Z48513 Z48512
	121366	280401_1 AI743515 AA405617 AW276706
	114609	116777_1 AA079505 AA079537
	115272	172113_1 AW015947 AA211890 AA279425
45	108338	112186_1 AA070773 AA070774
	108434	114012_1 AA078699 AA078782 AA075788
	123802	genbank_AA620448 AA620448
	102310	NOT_FOUND_entrez_U33839 U33839
	102636	entrez_U67092 U67092
50	104776	genbank_AA026349 AA026349
	120504	genbank_AA256837 AA256837
	113502	genbank_T89130T89130
	108499	genbank_AA083103 AA083103
	101308	entrez_L41390 L41390
55	108629	genbank_AA102425 AA102425
	103098	221_215 M86361 Z26593 X02850 D13070 AE000659 M17649 M87869 M87871 X61077 M16285 AF018169 X61079 S59351 X60142 AF043169
	103241	entrez_X76223 X76223
	103508	entrez_Y10141 Y10141
	103575	entrez_Z26256 Z26256
60	119514	NOT_FOUND_entrez_W37937 W37937
	121082	genbank_AA398722 AA398722
	128634	AA464918_at AA464918
	105817	genbank_AA397825 AA397825
	121518	genbank_AA412155 AA412155
65	114449	genbank_AA020736 AA020736
	114648	genbank_AA101056 AA101056
	121950	genbank_AA429515 AA429515
	107723	genbank_AA015957 AA015957

Table 3A shows 452 genes up-regulated in chronically diseased lung relative to normal lung. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 59680 probesets on the Eos/Alfymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number				
	ExAccn:	Exemplar Accession number, Genbank accession number				
	UnigeneID:	Unigene number				
	Unigene Title:	Unigene gene title				
10	R1:	80th percentile of AI for chronically diseased lung samples divided by the 90th percentile of AI for normal lung samples.				
	R2:	80th percentile of AI for chronically diseased lung samples divided by the 90th percentile of normal lung samples, squamous cell carcinomas and adenocarcinomas				
15	R3:	70th percentile of AI for chronically diseased lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of normal lung samples, squamous cell carcinomas and adenocarcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples				
	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
20	135423	U50531	Hs.138751	Human BRCA2 region, mRNA sequence CG030	12.40	
	135378	AW961818	Hs.24379	MUM2 protein		2.13
	135346	NM_000928	Hs.992	phospholipase A2, group IB (pancreas)		
	135235	AW298244	Hs.293507	ESTs	12.40	
	135057	U90268	Hs.93810	cerebral cavernous malformations 1	11.67	
25	134951	BE305081	Hs.169358	hypothetical protein		8.00
	134799	M36821	Hs.89690	GRO3 oncogene		8.20
	134786	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous		
	134772	NM_000829	Hs.163697	glutamate receptor, ionotropic, AMPA 4	29.80	
	134752	BE246762	Hs.89499	arachidonate 5-lipoxygenase		1.93
	134749	T28499	Hs.89485	carbonic anhydrase IV		2.07
30	134696	BE326276	Hs.8861	ESTs		
	134636	NM_005582	Hs.87205	lymphocyte antigen 64 (mouse) homolog, r	13.60	
	134627	A018768	Hs.12482	glyceronephosphate O-acyltransferase		1.92
	134622	AW975159	Hs.293097	ESTs, Weakly similar to A55380 facio-gen		1.92
	134570	U66615	Hs.172280	SW/SNF related, matrix associated, acti	13.20	
35	134561	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h		1.78
	134468	NM_001772	Hs.83731	CD33 antigen (gp67)		
	134417	NM_006416	Hs.82921	solute carrier family 35 (CMP-sialic aci		6.20
	134343	D50683	Hs.82028	transforming growth factor, beta recepto		
	134323	BE170651	Hs.8700	deleted in liver cancer 1		
40	134300	NM_001430	Hs.8136	endothelial PAS domain protein 1		
	134299	AW580939	Hs.97199	complement component C1q receptor		
	134253	X52075	Hs.80738	sialophorin (gpL115, leukosialin, CD43)	20.60	
	134182	D52059	Hs.7972	KIAA0871 protein	12.20	
	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec		
45	133978	AF035718	Hs.78061	transcription factor 21		
	133835	AI677897	Hs.76640	RGC32 protein		
	133651	AI301740	Hs.173381	dihydropyrimidinase-like 2		
	133633	D21262	Hs.75337	nucleolar and colloid-body phosphoprotein	15.20	
	133565	AW955776	Hs.313500	ESTs, Moderately similar to ALU7_HUMAN A		
50	133548	AW946384	Hs.178112	DNA segment, single copy probe LNS-CAUL		1.77
	133488	AA335295	Hs.74120	adipose specific 2		
	133478	X83703	Hs.31432	cardiac ankyrin repeat protein		2.08
	133337	AF085983	Hs.293676	ESTs		9.60
	133200	AB037715	Hs.183639	hypothetical protein FLJ10210		1.77
55	133153	AF070592	Hs.66170	HSKM-B protein	30.60	
	133130	AI128606	Hs.6557	zinc finger protein 161	22.60	
	133120	NM_003278	Hs.65424	tetranectin (plasminogen-binding protein		
	132928	AW168082	Hs.169449	protein kinase C, alpha	13.80	
	132836	AB023177	Hs.29900	KIAA0960 protein		
60	132799	W73311	Hs.169407	SAC2 (suppressor of actin mutations 2,	41.60	
	132742	AA025480	Hs.292812	ESTs, Weakly similar to T33468 hypotheti	40.40	
	132548	X12830	Hs.193400	interleukin 6 receptor		7.20
	132476	AL119844	Hs.49476	Homo sapiens clone TUA8 Crl-du-chat regi		4.76
	132439	AK001942	Hs.4863	hypothetical protein DKFP566A1524		1.88
65	132240	AB018324	Hs.42676	KIAA0781 protein	21.20	
	132210	NM_007203	Hs.42322	A kinase (PRKA) anchor protein 2		1.99
	132199	AL041299	Hs.165084	ESTs	15.20	
	131751	T96555	Hs.31562	ESTs		1.76
	131745	AI828559	Hs.31447	ESTs, Moderately similar to A46010 X-II	27.80	
70	131694	NM_000246	Hs.3076	MHC class II transactivator		4.00
	131686	NM_012296	Hs.30687	GRB2-associated binding protein 2		
	131676	AI126821	Hs.30514	ESTs		6.20
	131629	Z45794	Hs.238809	ESTs	21.40	
	131589	C18825	Hs.29191	epithelial membrane protein 2		
75	131536	AA019201	Hs.269210	ESTs		9.40
	131517	AB037789	Hs.263395	sema domain, transmembrane domain (TM),		3.59
	131355	R52804	Hs.25956	DKFP564D206 protein		4.48
	131253	R71802	Hs.24853	ESTs	15.00	
	131207	AF104266	Hs.24212	latrophilin		1.75
80	131156	AI472209	Hs.323117	ESTs		1.84
	131066	AW169287	Hs.22588	ESTs		3.54
	131061	NS4328	Hs.268744	KIAA1796 protein		
	131053	AA348541	Hs.296261	guanine nucleotide binding protein (G pr		1.93
	130895	AA641767	Hs.21015	hypothetical protein DKFP564L0864 simil	16.60	
85	130762	D84371	Hs.1898	paraoxonase 1	12.00	

	130657	AW337575	Hs.201591	ESTs		
	130655	AI831952	Hs.17409	cysteine-rich protein 1 (intestinal)		
	130589	AL110226	Hs.16441	DKFZP434H204 protein		2.08
5	130562	D50402	Hs.182611	solute carrier family 11 (proton-coupled)		1.91
	130555	R69743	Hs.116774	integrin, alpha 1	9.60	
	130365	W56119	Hs.155103	eukaryotic translation initiation factor	11.60	
	130273	AW972422	Hs.153853	MAD (mothers against decapentaplegic, Dr	6.60	
	130259	NM_000328	Hs.153514	retinitis pigmentosa GTPase regulator		1.91
10	130090	H97878	Hs.132390	zinc finger protein 36 (KOX 18)	21.20	
	129958	R27496	Hs.1378	annexin A3	5.05	
	129898	AI672731	Hs.13256	ESTs		
	129875	AA181018	Hs.13056	hypothetical protein FLJ13920	18.60	
	129699	AB007899	Hs.12017	homolog of yeast ubiquitin-protein ligase		
	129626	F13272	Hs.111334	ferritin, light polypeptide		
15	129598	N30436	Hs.11556	Homo sapiens cDNA FLJ12566 fis, clone NT	22.63	
	129593	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f		
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1		2.53
	129527	AA769221	Hs.270847	delta-tubulin	39.20	
20	129402	W72062	Hs.11112	ESTs		2.11
	129385	AA172106	Hs.110950	Rag C protein	15.20	
	129315	NM_014563	Hs.174038	spondyloepiphyseal dysplasia, late	12.40	
	129312	T97579	Hs.110334	ESTs, Weakly similar to I78885 serine/th	20.83	
	129240	AA361258	Hs.237868	interleukin 7 receptor		1.95
	129210	AL039940	Hs.202949	KIAA1102 protein		
25	129122	AW958473	Hs.301957	nudix (nucleoside diphosphate linked mol	4.20	
	129057	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)		
	128946	Y13153	Hs.107318	kynurenine 3-monooxygenase (kynurenine 3	5.20	
	128798	AF015525	Hs.302043	chemokine (C-C motif) receptor-like 2		
30	128789	AW368576	Hs.139851	caveolin 2		2.24
	128778	AA504776	Hs.186709	ESTs, Weakly similar to I38022 hypothet	12.20	
	128766	AW160432	Hs.296450	craniofacial development protein 1	26.40	
	128631	R44238	Hs.155546	KIAA1080 protein; Golgi-associated, gamm		1.78
	128624	BE154765	Hs.102647	ESTs, Weakly similar to TRHY_HUMAN TRICH		2.51
35	128609	NM_003616	Hs.102456	survival of motor neuron protein interac	16.00	
	128603	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHIT	12.80	
	128598	AA305407	Hs.102308	potassium inwardly-rectifying channel, s	4.00	
	128458	H55864	Hs.56340	ESTs		
40	128061	AF150882	Hs.186877	sodium channel, voltage-gated, type XII,	17.20	
	127968	AA830201	Hs.124347	ESTs	21.30	
	127959	AI302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L		
	127944	AI557081	Hs.262476	S-adenosylmethionine decarboxylase 1	10.60	
	127925	AA805151	Hs.3628	mitogen-activated protein kinase kinase	13.40	
45	127896	AI669586	Hs.222194	ESTs	7.00	
	127859	AA761802	Hs.291559	ESTs	14.00	
	127817	AA836641	Hs.163085	ESTs	14.00	
	127742	AW293496	Hs.180138	ESTs	11.00	
	127628	AI240102	Hs.322430	NDRG family, member 4	11.10	
	127609	X80031	Hs.530	collagen, type IV, alpha 3 (Goodpasture		
50	127582	AA908954	Hs.130844	ESTs	19.60	
	127543	AK000787	Hs.157392	Homo sapiens cDNA FLJ20780 fis, clone CO	15.40	
	127535	AA568424	Hs.164450	ESTs	17.50	
	127404	AI379920	Hs.270224	ESTs	14.60	
	127396	L31968	Hs.187991	DKFZP564A122 protein	15.40	
55	127374	AA442797	Hs.312110	ESTs, Weakly similar to I38022 hypothet	14.60	
	127346	AA203616	Hs.44896	DnaJ (Hsp40) homolog, subfamily B, membe	21.00	
	127340	BE047653	Hs.119183	ESTs, Weakly similar to ZN91_HUMAN ZINC	15.80	
	127307	AW962712	Hs.126712	ESTs, Weakly similar to AF191020 1 E2IG5		
	127242	AW390395	Hs.181301	cathepsin S	22.60	
60	127167	AA625690	Hs.190272	ESTs	21.40	
	127046	AA321948	Hs.293968	ESTs	41.20	
	126928	AA480902	Hs.137401	ESTs	11.00	
	126900	AF137386	Hs.12701	plasmalogen		1.78
	126852	AA399961		gb:zu68c01.r1 Soares_testis_NHT Homo sap	5.60	
65	126816	AA248234		gb:csq2228.seq.F Human fetal heart, Lamb	12.20	
	126812	AB037860	Hs.173933	nuclear factor 1A	17.19	
	126666	AA648886	Hs.151999	ESTs	13.57	
	126645	AA316181	Hs.61635	sbx transmembrane epithelial antigen of	15.40	
	126592	AI611153	Hs.6093	Homo sapiens cDNA: FLJ22783 fis, clone K	4.67	
	126556	AF255303	Hs.112227	membrane-associated nucleic acid binding	18.00	
70	126433	AA325606		gb:EST28707 Cerebellum II Homo sapiens c	16.77	
	126299	AW979155	Hs.298275	amino acid transporter 2	14.60	
	126218	AL049801	Hs.13649	Novel human gene mapping to chromosome 13	3.50	
	126182	AA721331	Hs.293771	ESTs	13.40	
75	126177	AW752782	Hs.129750	hypothetical protein FLJ10546	18.20	
	126142	H86261	Hs.40568	ESTs	14.00	
	126077	M78772	Hs.210836	ESTs	16.59	
	125994	AI990529	Hs.270799	ESTs	17.40	
	125934	AA193325	Hs.32646	hypothetical protein FLJ21901	13.00	
80	125847	AW161885	Hs.249034	ESTs	49.57	
	125831	H04043		gb:bj45c03.r1 Soares placenta Nb2HP Homo		
	125731	R61771	Hs.26912	ESTs	13.20	
	125676	BE612918	Hs.151973	hypothetical protein FLJ23511	11.20	
	125661	F18572	Hs.22978	ESTs, Weakly similar to ALU4_HUMAN ALU S		
85	125552	H09701	Hs.278366	ESTs, Weakly similar to I38022 hypotheti	12.60	
	125489	H49193	Hs.124984	ESTs, Moderately similar to ALU7_HUMAN A	33.40	

	125422	AA903229	Hs.153717	ESTs		1.80
	125331	AA22996	Hs.161378	ESTs	38.00	
	125309	T12411	Hs.183745	hypothetical protein FLJ13456	18.20	
5	125167	AL137540	Hs.102541	netrin 4		1.95
	125139	AW194933	Hs.9788	hypothetical protein MGC10924 similar to		1.84
	125042	T78906	Hs.269432	ESTs, Moderately similar to ALU1_HUMAN	21.80	
	124711	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	10.60	
	124631	NM_014053	Hs.270594	FLVCR protein	23.20	
	124578	N58321	Hs.231500	EST	21.43	
10	124574	AL036596	Hs.42322	A kinase (PKA) anchor protein 2		1.77
	124472	N52517	Hs.102670	EST	37.20	
	124438	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A		
	124357	N22401		gb:yyw37g07.s1 Morton Fetal Cochlea Homo	14.64	
	124306	AW973078	Hs.293039	ESTs	4.00	
15	124214	H56608	Hs.151323	ESTs		
	124097	AW298235	Hs.101689	ESTs	27.20	
	123978	T89832	Hs.170278	ESTs		2.03
	123972	T46848	Hs.70337	immunoglobulin superfamily, member 4	6.00	
	123961	AL050184	Hs.21610	DKFZP434B203 protein		1.79
20	123936	NM_004673	Hs.241519	angiopoietin-like 1	15.80	
	123902	AA620448		gb:ae58c09.s1 Stratagene lung carcinoma	4.23	
	123734	AA609861	Hs.312447	ESTs	4.20	
	123619	AA602964		gb:nc97c02.s1 NCL_CGAP_Pr2 Homo sapiens	33.60	
	123596	AA421130	Hs.112640	EST	10.93	
25	123476	AA384564	Hs.108829	ESTs		2.18
	123340	AA504264	Hs.182937	peptidylprolyl isomerase A (cyclophilin	11.20	
	123190	AA489212	Hs.105228	EST	14.20	
	123136	AW451999	Hs.194024	ESTs		7.00
	123073	AA485061	Hs.105652	ESTs	31.20	
30	123055	AA482005	Hs.105102	ESTs, Weakly similar to reverse transcri		4.80
	122699	AA456130	Hs.301721	KIAA1255 protein		5.00
	122679	AA811286	Hs.192837	ESTs, Weakly similar to ALU5_HUMAN ALU S	14.40	
	122633	NM_001546	Hs.34853	inhibitor of DNA binding 4, dominant neg		
	122553	AA451884	Hs.190121	ESTs	40.00	
35	122544	AW973253	Hs.292689	ESTs	15.40	
	122485	AA524547	Hs.160318	FXFD domain-containing ion transport reg		1.81
	122211	AA300900	Hs.98849	ESTs, Moderately similar to AF161511.1 H	12.10	
	122127	AW207175	Hs.106771	ESTs		1.95
	122011	AA431082		gb:zw78a10.s1 Soares_testis_NHT Homo sap		1.89
40	121992	AI860775	Hs.98506	ESTs	3.60	
	121989	W56487	Hs.193784	Homo sapiens mRNA; cDNA DKFZp586K1922 (f		2.01
	121835	AB033030	Hs.300570	KIAA1204 protein		1.85
	121726	AF241254	Hs.178098	angiotensin I converting enzyme (peptidy	12.43	
45	121690	AV660305	Hs.110288	ESTs		1.82
	121643	AA640987	Hs.193767	ESTs		
	121633	AA417011	Hs.98175	EST	14.00	
	121622	AA416931	Hs.126065	ESTs	16.40	
	121497	AA412031	Hs.97901	EST	11.20	
	121351	AW206227	Hs.287727	hypothetical protein FLJ23132	12.20	
50	121314	W07343	Hs.182538	phospholipid scramblase 4		1.83
	121242	AA400857	Hs.97509	ESTs	22.40	
	121059	AA332883		gb:z174e03.r1 Soares_testis_NHT Homo sap	14.80	
	120934	AA226198		gb:nc26a07.s1 NCL_CGAP_Pr1 Homo sapiens	21.20	
55	120755	AA312934	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone		1.79
	120637	AA811804		gb:ob39a05.s1 NCL_CGAP_GC81 Homo sapiens	20.00	
	120484	AA253170	Hs.96473	EST	40.20	
	120336	N85785	Hs.181165	eukaryotic translation elongation factor	6.60	
	120266	AI807264	Hs.205442	ESTs, Weakly similar to T34036 hypotheti	16.80	
	120132	W57554	Hs.125019	ESTs	4.73	
60	120041	AA830882	Hs.59368	ESTs		1.75
	119996	W88996	Hs.59134	EST	7.20	
	119970	AA757718	Hs.93581	hypothetical protein FLJ10512	11.20	
	119861	W78816	Hs.49943	ESTs, Weakly similar to S65657 alpha-1C-		3.78
	119824	W74536	Hs.184	advanced glycosylation end product-speci		
65	119740	AW021407	Hs.21058	hypothetical protein	20.20	
	119271	AI061118	Hs.65328	Fanconi anemia, complementation group F	15.20	
	119221	C14322	Hs.250700	trypsin beta 1		
	119126	R45175	Hs.117183	ESTs	12.60	
	119073	BE245360	Hs.279477	ESTs		
70	118928	AA312799	Hs.283689	activator of CREM in testis	10.00	
	118901	AW292577	Hs.94445	ESTs	3.96	
	118661	AL137554	Hs.49927	protein kinase NYD-SP15	9.60	
	118607	AI377444	Hs.54245	ESTs, Weakly similar to S65824 reverse t	10.40	
	118449	AI813865	Hs.164478	hypothetical protein FLJ21939 similar to		1.90
75	118416	N66028	Hs.49105	FKBP-associated protein	16.20	
	118379	N64491	Hs.48990	ESTs		4.00
	118329	N63520		gb:yy62f01.s1 Soares_multiple_sclerosis_	6.60	
	118320	N63451	Hs.141600	ESTs, Weakly similar to alternatively s	3.80	
	118253	AA497044	Hs.20887	hypothetical protein FLJ10392	17.60	
80	118124	N56968	Hs.46707	chromosome 21 open reading frame 37	14.00	
	118056	AB037746	Hs.42768	hypothetical protein DKFZp761O0113		1.86
	118032	N52802	Hs.47544	EST	5.00	
	117840	T26379	Hs.48802	Homo sapiens clone 23632 mRNA sequence	4.00	
	117404	N39725	Hs.15220	zinc finger protein 106		1.90
85	117314	N32498	Hs.42829	ESTs	14.20	

	117209	W03011	Hs.306881	MSTP043 protein			
	117023	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f	20.20	2.31	
	116814	H50834		gb:yp86a10.s1 Soares fetal liver spleen			
5	116784	AB007979	Hs.301281	Homo sapiens mRNA, chromosome 1 specific	3.51		
	116766	AI508657	Hs.95097	ESTs	16.20	6.80	
	116712	AW901618	Hs.61935	Homo sapiens mRNA; cDNA DKFZp7611071 (fr			
	116707	H10344	Hs.49050	ESTs, Weakly similar to A Chain A, Human	18.60		
	116351	AL133623	Hs.82501	similar to mouse Xrn1 / Dhm2 protein	19.40		
10	116279	AW971248	Hs.291289	ESTs, Weakly similar to ALU1_HUMAN ALU S			
	116166	AL039940	Hs.202949	KIAA1102 protein		2.13	
	116152	AL040521	Hs.15220	zinc finger protein 106		1.75	
	116117	BE613410	Hs.31575	SEC63, endoplasmic reticulum translocon	13.20		
	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	30.11		
15	115985	AA001732	Hs.173233	hypothetical protein FLJ10970		2.36	
	115955	AF263613	Hs.44198	Intracellular membrane-associated calciu	18.20		
	115844	AI373062	Hs.332938	hypothetical protein MGC5370	18.57		
	115683	AF255910	Hs.54650	junctional adhesion molecule 2	23.00		
	115673	AA406341	Hs.269308	Homo sapiens cDNA FLJ11991 fis, clone HE	11.82		
	115672	AI889110	Hs.73251	ESTs	10.60		
20	115566	AI142336	Hs.43977	Human DNA sequence from clone RP11-196N1		1.76	
	115313	AA808001	Hs.184411	albumin	25.20		
	115279	AW964897	Hs.290825	ESTs		8.00	
	115230	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L		1.80	
	115110	AK001671	Hs.11387	KIAA1453 protein	14.20		
25	114999	BE246481	Hs.87856	ESTs	19.20		
	114930	AA237022	Hs.188717	ESTs		5.60	
	114922	AA235672	Hs.87491	ESTs		3.60	
	114837	BE244930	Hs.166895	ESTs	43.70		
	114769	AA149060	Hs.296100	ESTs	11.00		
30	114761	AA143781	Hs.126280	hypothetical protein FLJ23393	14.00		
	114736	AI510347	Hs.103812	ESTs, Moderately similar to ALU1_HUMAN A		4.20	
	114596	AA310162	Hs.169248	cytochrome c	10.71		
	114518	AW163267	Hs.106469	suppressor of var1 (S.cerevisiae) 3-like	20.40		
	114455	H37908	Hs.271616	ESTs, Weakly similar to ALU8_HUMAN ALU S	20.40		
35	114452	AI369275	Hs.243010	Homo sapiens cDNA FLJ14445 fis, clone HE		17.20	
	114359	NM_016929	Hs.283021	chloride intracellular channel 5		2.09	
	114357	R41677	Hs.6107	Homo sapiens cDNA FLJ14839 fis, clone OV	12.40		
	114251	H15261	Hs.21948	ESTs		2.00	
40	114138	AW394793	Hs.15740	Homo sapiens mRNA; cDNA DKFZp434E033 (fr		11.40	
	114124	W57554	Hs.125019	ESTs		6.04	
	113946	AW083883	Hs.37896	Homo sapiens cDNA FLJ13510 fis, clone PL		1.82	
	113695	T96965	Hs.17948	ESTs, Weakly similar to ALU8_HUMAN IIII			
	113606	NM_013343	Hs.278951	NAG-7 protein		2.15	
	113590	R49642	Hs.142447	ESTs, Weakly similar to ALU1_HUMAN ALU S		3.60	
45	113560	T91015	Hs.268826	ESTs	32.00		
	113552	AI654223	Hs.16026	hypothetical protein FLJ23191			
	113540	AW152618	Hs.16757	ESTs			
	113502	T89130		gb:ye12d01.s1 Stratagene lung (937210) H		8.35	
50	113288	AI076838	Hs.12967	ESTs	12.40		
	113252	NM_004469	Hs.11392	c-fos induced growth factor (vascular en		4.27	
	113238	R45467	Hs.189813	ESTs			
	113203	AA743563	Hs.10305	ESTs	21.20		
	113195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom		1.92	
55	113089	T40707	Hs.270862	ESTs	14.33		
	113076	AF033199	Hs.8198	zinc finger protein 204		6.00	
	113009	T23699	Hs.7246	ESTs		9.40	
	112937	AI694320	Hs.6295	ESTs, Weakly similar to T17248 hypotheti		12.20	
	112891	T03927	Hs.293147	ESTs, Moderately similar to A46010 X-H	10.57		
60	112794	R97018		gb:yg74b08.s1 Soares fetal liver spleen	26.60		
	112691	R88708	Hs.220647	ESTs	15.33		
	112602	AW004045	Hs.203365	ESTs	15.60		
	112366	AF035318	Hs.12533	Homo sapiens clone 23705 mRNA sequence	15.40		
	112210	R49645	Hs.7004	ESTs	14.00		
65	112064	AL049390	Hs.22689	Homo sapiens mRNA; cDNA DKFZp586O1318 (f	13.00		
	111998	R42379	Hs.138283	ESTs	11.00		
	111987	NM_015310	Hs.6763	KIAA0942 protein.	22.40		
	111803	AA593731	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A		1.77	
	111737	H04607	Hs.9218	ESTs		1.86	
70	111605	T91061	Hs.194178	ESTs, Moderately similar to PC4259 ferri	23.00		
	111510	R07856	Hs.16355	ESTs	11.02		
	111341	AL157484	Hs.22483	Homo sapiens mRNA; cDNA DKFZp762M127 (fr		1.88	
	111280	AA373527	Hs.19385	CGI-58 protein	18.40		
	111247	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp56482062 (f			
75	111232	AI247763	Hs.16928	ESTs	27.60		
	110942	R63503	Hs.28419	ESTs	14.80		
	110924	AW058463	Hs.12940	zinc-fingers and homeoboxes 1	24.71		
	110837	H03109	Hs.108920	HT018 protein		2.18	
	110824	AI767183	Hs.26942	ESTs	12.20		
80	110776	AB032417	Hs.19545	frizzled (Drosophila) homolog 4		1.75	
	110576	H60869	Hs.37889	ESTs	13.00		
	110369	AK000768	Hs.107872	hypothetical protein FLJ20761		5.60	
	110099	R44557	Hs.23748	ESTs		2.31	
	109984	AI796320	Hs.10299	Homo sapiens cDNA FLJ13545 fis, clone PL			
	109958	AA001266	Hs.133521	ESTs	11.25		
85	109893	AA884208	Hs.30484	ESTs		2.68	

	109842	AW818436	Hs.23590	solute carrier family 16 (monocarboxylic	23.83		
	109837	H00656	Hs.29792	ESTs, Weakly similar to I38022 hypothe		3.91	
	109796	AJ800515	Hs.12024	ESTs	17.20		
	109688	R41900	Hs.22245	ESTs	9.60		
5	109648	H17800	Hs.7154	ESTs	22.80		
	109613	H47315	Hs.27519	ESTs			
	109550	AW021488	Hs.26981	ESTs			
	109523	AW193342	Hs.24144	ESTs			
	109472	AK001989	Hs.91165	hypothetical protein	6.00	1.89	
10	109355	AA524525	Hs.48297	DKFZP586C1620 protein	15.00		
	109260	AW978515	Hs.131915	KIAA0863 protein	25.60		
	108781	AA128554		gbzn98g07.s1 Stratagene fetal retina 93	14.20		
	108663	BE219231	Hs.292653	ESTs, Weakly similar to T26845 hypothe	11.00		
	108573	AA086005		gbzl84c04.s1 Stratagene colon (937204)	26.00		
15	108480	AL133092	Hs.68055	hypothetical protein DKFZp434J0428			
	108382	NM_006770	Hs.67726	macrophage receptor with collagenous str		1.83	
	108174	AA055632	Hs.303070	ESTs	15.20		
	108138	AL049990	Hs.51515	Homo sapiens mRNA; cDNA DKFZp554G112 (fr		3.60	
	108087	AA045708	Hs.40545	ESTs	15.44		
20	108048	AJ797341	Hs.165195	Homo sapiens cDNA FLJ14237 fis, clone NT		11.40	
	108041	AW204712	Hs.61957	ESTs			
	107997	AL049176	Hs.82223	chordin-like	4.76		
	107994	AA036811	Hs.48469	LIM domains containing 1			
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	14.20		
25	107681	BE379594	Hs.49136	ESTs, Moderately similar to ALU7_HUMAN A	51.80		
	107666	AA010611	Hs.60418	EST	29.20		
	107332	T87750	Hs.183297	DKFZP566F2124 protein	10.73		
	107292	BE166479	Hs.4789	Homo sapiens serologically defined breas	32.00		
30	107230	AJ034467	Hs.34650	ESTs	17.40		
	107168	W57578	Hs.237955	RAB7, member RAS oncogene family	10.43		
	107160	AA314490	Hs.27669	KIAA1563 protein	11.40		
	107054	AJ076459	Hs.15978	KIAA1272 protein			
	107029	AF264750	Hs.288971	myeloid/lymphoid or mixed-lineage leukem	21.40		
35	106999	H93281	Hs.10710	hypothetical protein FLJ20417	35.80		
	106954	AF128847	Hs.204038	indolethylamine N-methyltransferase		1.76	
	106870	AJ937330	Hs.26530	serum deprivation response (phosphatidyl			
	106865	AW192535	Hs.19479	ESTs	13.40		
	106844	AA485055	Hs.158213	sperm associated antigen 6	7.13		
40	106820	NM_016831	Hs.12592	period (Drosophila) homolog 3	7.00		
	106818	AK002135	Hs.3542	hypothetical protein FLJ11273	13.00		
	106797	AJ768801	Hs.169943	Homo sapiens cDNA FLJ13569 fis, clone PL		2.05	
	106773	AA478109	Hs.188833	ESTs			
	106747	NM_007118	Hs.171957	triple functional domain (PTPRF interact	12.60		
45	106743	BE613328	Hs.21938	hypothetical protein FLJ12492	10.60		
	106667	AW360847	Hs.16578	ESTs			
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr		2.40	
	106567	AW450408	Hs.86412	chromosome 9 open reading frame 5		1.78	
	106562	AL031846	Hs.152151	plakophilin 4		1.76	
50	106536	AA329648	Hs.23804	ESTs, Weakly similar to PN0099 son3 prot		2.19	
	106533	AL134708	Hs.145998	ESTs	23.20		
	106507	AA259068	Hs.267819	protein phosphatase 1, regulatory (inhib	15.20		
	106490	AA404265	Hs.115537	putative dipeptidase			
	106474	BE383668	Hs.42484	hypothetical protein FLJ10618	10.44		
55	106211	AA428240	Hs.126083	ESTs		29.80	
	105986	AB037722	Hs.8707	KIAA1301 protein	3.70		
	105894	AJ904740	Hs.25691	receptor (calcitonin) activity modifying		1.94	
	105847	AW984490	Hs.32241	ESTs, Weakly similar to S65657 alpha-1C-		1.75	
	105803	AW747996	Hs.160999	ESTs, Moderately similar to A56194 throm		2.47	
60	105731	AA834664	Hs.29131	nuclear receptor coactivator 2	10.71		
	105729	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds			
	105688	AJ299139	Hs.17517	ESTs	23.40		
	105510	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds	37.20		
	105101	H63202	Hs.38163	ESTs	8.30		
65	104989	R65998	Hs.285243	hypothetical protein FLJ22029	8.09		
	104986	AW088826	Hs.117176	poly(A)-binding protein, nuclear 1		1.92	
	104969	AJ670947	Hs.78406	phosphatidylinositol-4-phosphate 5-kinas	5.40		
	104903	AA363223	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,	7.60		
	104896	AW015318	Hs.23165	ESTs	13.80		
70	104865	T79340	Hs.22575	Homo sapiens cDNA: FLJ21042 fis, clone C			
	104825	AA035613	Hs.141883	ESTs		1.87	
	104781	AA099904	Hs.21610	DKFZP434B203 protein		1.93	
	104776	AA026349		gbzj99f01.s1 Soares_pregnant_uterus_NbH	10.20		
	104691	U29690	Hs.37744	Homo sapiens beta-1 adrenergic receptor	5.69		
	104667	AJ239923	Hs.30098	ESTs	3.82		
75	104404	H58762		gb:EST00057 HE6W Homo sapiens cDNA clone	4.20		
	104392	AA076049	Hs.274415	Homo sapiens cDNA FLJ10229 fis, clone HE	27.20		
	104212	AB002298	Hs.173035	KIAA0300 protein		1.91	
	104074	AL162039	Hs.31422	Homo sapiens mRNA; cDNA DKFZp434M229 (fr	11.20		
	103749	AL135301	Hs.8768	hypothetical protein FLJ10849	10.86		
80	103645	AW246253	Hs.7043	succinate-CoA ligase, GDP-forming, alpha	12.00		
	103554	AJ878826	Hs.323469	caveolin 1, caveolae protein, 22kD		1.80	
	103541	AJ815601	Hs.79197	CD83 antigen (activated B lymphocytes, i			
	103496	Y09267	Hs.132821	flavin containing monooxygenase 2			
	103428	BE383507	Hs.78921	A kinase (PRKA) anchor protein 1	11.20		
85	103353	X89399	Hs.119274	RAS p21 protein activator (GTPase activa	19.80		

5	103295	X81479	Hs.2375	egf-like module containing, mucin-like,	3.60	
	103280	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula		1.76
	103100	NM_005574	Hs.184585	LIM domain only 2 (rhombotin-like 1)		2.15
	103025	NM_002837	Hs.123641	protein tyrosine phosphatase, receptor t		
	102698	M18657	Hs.1867	progastricin (pepsinogen C)		
10	102659	BE245169	Hs.211610	CUG triplet repeat, RNA-binding protein	11.00	
	102580	U60308	Hs.152981	CDP-diacylglycerol synthase (phosphatida	25.40	
	102417	AA034127	Hs.153487	signal transducing adaptor molecule (SH3	14.00	
	102363	NM_003734	Hs.198241	amine oxidase, copper containing 3 (vasc		
	102302	AA306342	Hs.69171	protein kinase C-like 2	10.86	
15	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11		
	102188	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	7.40	
	102151	T27013	Hs.3132	steroidogenic acute regulatory protein	16.40	
	101957	L28824	Hs.74101	spleen tyrosine kinase	15.40	
	101842	M93221	Hs.75182	mannose receptor, C type 1		
20	101771	NM_002432	Hs.153837	myeloid cell nuclear differentiation ant		
	101764	A198550	Hs.81256	S100 calcium-binding protein A4 (calcium		1.78
	101716	AF050658	Hs.2563	tachykinin, precursor 1 (substance K, su	18.80	
	101678	M62505	Hs.2161	complement component 5 receptor 1 (C5a)		2.22
	101447	M21305		gb:human alpha satellite and satellite 3	504.80	
25	101383	NM_000132	Hs.79345	coagulation factor VIII, procoagulant co	31.00	
	101346	A1738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N		1.75
	101345	NM_005795	Hs.152175	calcitonin receptor-like		
	101336	NM_006732	Hs.75678	FBJ murine osteosarcoma viral oncogene h		2.24
	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do		
30	101277	BE297626	Hs.296049	microfibrillar-associated protein 4		
	101262	L35854		gb:Human dystrophin (dp140) mRNA, 5' end	19.00	
	101168	NM_005308	Hs.211569	G protein-coupled receptor kinase 5		2.01
	101102	NM_003243	Hs.79059	transforming growth factor, beta recepto		
	101088	X70697	Hs.553	solute carrier family 6 (neurotransmitte	7.52	
35	101066	AW970254	Hs.889	Charot-Leyden crystal protein	19.38	
	100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte		1.91
	100893	BE245294	Hs.180789	S164 protein	15.40	
	100770	W25797.comp	Hs.177486	amyloid beta (A4) precursor protein (pro	11.20	
	100716	X89887	Hs.172350	HIR (histone cell cycle regulation defec	14.80	
40	100555	M69181		gb:Human nonmuscle myosin heavy chain-B	33.00	
	100425	NM_014747	Hs.78748	KIAA0237 gene product	16.20	
	100408	D86640	Hs.56045	src homology three (SH3) and cysteine ri	4.00	
	100382	D83407	Hs.156007	Down syndrome critical region gene 1-lik	4.24	
	100351	D64158			6.20	
45	100299	D49493	Hs.2171	growth differentiation factor 10	21.20	
	100134	AA305746	Hs.49	macrophage scavenger receptor 1		1.79
	100108	U09577	Hs.76873	hyaluronoglucosaminidase 2		
	100095	Z97171	Hs.78454	myocilin, trabecular meshwork inducible	5.40	
	100066				11.29	

TABLE 3B shows the accession numbers for those primekeys lacking unigenelD's for Table 3A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

55 Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
60	123619	371681_1 AA602964 AA609200
	126433	127143_1 AA325606 AA099517 N89423
	125831	1522905_1 H04043 D60988 D60337
	126816	122973_1 AA248234 AA090985
	126852	136135_1 AA399961 AA128347
65	121059	273450_1 AA393283 AA398528
	120637	200885_1 AA811804 AA809404 AA286907 AW977624
	122011	7617_2 AA431082
	120934	177521_1 AA226198 AA226513 AA383773
	123802	genbank_AA620448 AA620448
70	116814	genbank_H50834 H50834
	118329	genbank_N63520 N63520
	104404	H58762_at H58762
	104776	genbank_AA026349 AA026349
	113502	genbank_T89130T89130
75	101262	entrez_L35854 L35854
	108573	genbank_AA086005 AA086005
	101447	entrez_M21305 M21305
	124357	genbank_N22401 N22401
	108781	genbank_AA128654 AA128654
80	112794	genbank_R97018 R97018
	100351	entrez_D64158 D64158
	100555	figr_HT2245 M69181 M81105 U51039
85		

Table 4A shows 202 genes up-regulated in samples from patients treated with chemotherapy or radiotherapy. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigenelD:	Unigene number			
	Unigene Title:	Unigene gene title			
	R1:	average of AI for samples from patients treated with chemotherapy or radiotherapy divided by the average of AI for normal lung samples.			
10	Pkey	ExAccn	UnigenelD	Unigene Title	R1
	100113	NM_001269	Hs.84746	chromosome condensation 1	27.20
	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3	20.60
15	100210	D26361	Hs.3104	KIAA0042 gene product	20.40
	100225	D28539	Hs.167185	glutamate receptor, metabotropic 5	20.60
	100269	NM_001949	Hs.1189	E2F transcription factor 3	29.40
	100438	AA013051	Hs.91417	topoisomerase (DNA) II binding protein	23.50
	100877	X80821	Hs.27973	KIAA0874 protein	35.56
20	100893	BE245294	Hs.180789	S164 protein	43.40
	101273	Z11933	Hs.182505	POU domain, class 3, transcription facto	21.80
	101447	M21305		gbtHuman alpha satellite and satellite 3	193.60
	101649	AW959908	Hs.1690	heparin-binding growth factor binding pr	38.40
	101724	L11690	Hs.620	bulbosus pemphigoid antigen 1 (230/240kD)	198.80
25	101748	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	78.60
	101809	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	162.20
	101879	AA176374	Hs.243886	nuclear autoantigenic sperm protein (his	50.00
	101915	AF207881	Hs.155185	cytosolic ovarian carcinoma antigen 1	26.00
	101973	U41514	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp	37.20
30	102025	U04045	Hs.78934	mutS (E. coli) homolog 2 (colon cancer,	
	102031	U04898	Hs.2156	RAR-related orphan receptor A	32.00
	102052	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	51.20
	102391	AA296874	Hs.77494	deoxyguanosine kinase	13.90
	102420	U44060	Hs.14427	Homo sapiens cDNA: FLJ21800 fs, clone H	28.80
35	102610	U65011	Hs.30743	preferentially expressed antigen in mela	110.60
	102829	NM_005183	Hs.80962	neurotensin	116.80
	103000	NM_001975	Hs.146580	enolase 2, (gamma, neuronal)	2.30
	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	181.40
	103507	AJ000512	Hs.296323	serum/glucocorticoid regulated kinase	49.20
40	103587	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	86.60
	104660	BE298665	Hs.14846	Homo sapiens mRNA: cDNA DKFZp564D016 (fr	42.60
	104896	AW015318	Hs.23165	ESTs	29.40
	105038	AW503733	Hs.9414	KIAA1488 protein	21.50
	105298	BE387790	Hs.26369	hypothetical protein FLJ20287	32.80
45	105510	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds	20.20
	105667	AA767526	Hs.22030	paired box gene 5 (B-cell lineage specif	28.40
	106073	AL157441	Hs.17834	downstream neighbor of SON	25.40
	106205	AW965058	Hs.111583	ESTs, Weakly similar to t38022 hypotheti	32.00
	106516	AL137311	Hs.234074	Homo sapiens mRNA: cDNA DKFZp761G02121 (40.60
50	106533	AL134708	Hs.145998	ESTs	59.80
	106575	AW970602	Hs.105421	ESTs	43.40
	106654	AW075485	Hs.286049	phosphoserine aminotransferase	50.80
	106851	AA58623		gbt:tk04g09.x1 NCI_CGAP_Lu24 Homo sapiens	53.40
	106995	AB023139	Hs.37892	KIAA0922 protein	20.88
55	107332	T87750	Hs.183297	DKFZP566F2124 protein	23.60
	107532	AA443473	Hs.173684	Homo sapiens mRNA: cDNA DKFZp762G207 (fr	57.20
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	49.00
	108609	BE409857	Hs.69499	hypothetical protein	19.67
	108780	AU076442	Hs.117938	collagen, type XVII, alpha 1	48.17
60	109166	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabkines	59.20
	109260	AW978515	Hs.131915	KIAA0863 protein	28.60
	109280	AK001355	Hs.279610	hypothetical protein FLJ10493	22.80
	109292	AW975746	Hs.188662	KIAA1702 protein	
	109384	AA219172	Hs.86849	ESTs	21.00
65	109415	U80736	Hs.110826	trinucleotide repeat containing 9	31.60
	109445	AA232103	Hs.189915	ESTs	24.20
	109502	AW967069	Hs.211556	hypothetical protein MGC5487	21.40
	109633	AW003785	Hs.170267	ESTs	20.40
	109786	AI989482	Hs.146286	kinesin family member 13A	19.60
70	109958	AA001266	Hs.133521	ESTs	24.00
	110920	N47224	Hs.20521	HMT1 (hnRNP methyltransferase, S. cerevi	28.40
	110924	AW058463	Hs.12940	zinc-fingers and homeoboxes 1	36.00
	111084	H44186	Hs.15456	PDZ domain containing 1	61.20
	111132	AB037807	Hs.83293	hypothetical protein	24.60
75	111229	AW389845	Hs.110855	ESTs	27.20
	111337	AA837396	Hs.263925	LIS1-interacting protein NUDE1, rat homo	48.00
	111987	NM_015310	Hs.6763	KIAA0942 protein	37.80
	112046	AA383343	Hs.22116	CDC14 (cell division cycle 14, S. cerevi	26.80
	112268	W39609	Hs.22003	solute carrier family 6 (neurotransmitte	63.80
80	112685	R87650	Hs.33439	ESTs, Weakly similar to ALU1_HUMAN ALU	26.40
	112871	AL110216	Hs.12285	ESTs, Weakly similar to I55214 salivary	47.64
	112897	AW206453	Hs.3782	ESTs	22.00
	112973	AB033023	Hs.318127	hypothetical protein FLJ10201	65.00
	112992	AL157425	Hs.133315	Homo sapiens mRNA: cDNA DKFZp761J1324 (f	42.00
85	113073	N39342	Hs.103042	microtubule-associated protein 1B	55.40

	113494	T91451	Hs.86538	ESTs	22.80
	113560	T91015	Hs.268626	ESTs	22.80
	113849	AA457211	Hs.8858	bromodomain adjacent to zinc finger doma	51.80
	113950	AI267652	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (fr	28.20
5	114339	AA782845	Hs.22790	ESTs	20.20
	114365	H42169	Hs.18653	hypothetical protein FLJ14627	21.00
	114455	H37908	Hs.271616	ESTs, Weakly similar to ALU8_HUMAN ALU S	25.80
	114518	AW163267	Hs.106469	suppressor of var1 (S.cerevisiae) 3-like	23.60
10	114824	AA960961	Hs.305953	zinc finger protein 63 (HPF1)	27.20
	114837	BE244930	Hs.166895	ESTs	30.20
	114974	AW966931	Hs.179652	nucleosome assembly protein 1-like 1	20.80
	115075	AA814043	Hs.88045	ESTs	30.60
	115084	BE383668	Hs.42484	hypothetical protein FLJ10618	28.86
	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	38.00
15	115313	AA808001	Hs.184411	albumin	22.60
	115697	D31382	Hs.63325	transmembrane protease, serine 4	173.60
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	27.77
	116090	AI591147	Hs.61232	ESTs	20.80
	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	164.20
20	116399	AA889120	Hs.110637	homeo box A10	38.00
	117099	H93699		gb:vv16a11.s1 Soares fetal liver spleen	21.60
	117881	AF161470	Hs.260622	butyrate-induced transcript 1	49.40
	118091	AW005054	Hs.47883	ESTs, Weakly similar to KCC1_HUMAN CALCI	22.40
25	118138	AA374756	Hs.93560	Homo sapiens mRNA for KIAA1771 protein,	22.00
	118720	N73515		gb:za49d07.s1 Soares fetal liver spleen	20.00
	118873	AI824009	Hs.44577	ESTs	19.40
	119126	R45175	Hs.117183	ESTs	111.20
	119717	AA918317	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro	33.00
30	119940	AL050097	Hs.272531	DKFZP586B0319 protein	31.00
	120266	AI807264	Hs.205442	ESTs, Weakly similar to T34036 hypotheti	20.20
	120615	AA258356		gb:cz59c10.s1 Soares_NhHMPu_S1 Homo sapi	25.00
	120859	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol	95.40
	120983	AA398209	Hs.97587	EST	105.20
35	121054	AW976570	Hs.97387	ESTs	38.80
	121369	AW450737	Hs.128791	CGI-09 protein	41.60
	122335	AA443258	Hs.241551	chloride channel, calcium activated, fam	30.80
	122612	AA974832	Hs.128708	ESTs	19.60
	123130	AA487200		gb:ab19f02.s1 Stratagene lung (937210) H	33.20
40	123440	AI733692	Hs.112488	ESTs	23.17
	123596	AA421130	Hs.112640	EST	23.00
	123619	AA602964		gb:nc97c02.s1 NCI_CGAP_Pr2 Homo sapiens	28.80
	124006	AI147155	Hs.270016	ESTs	77.60
	124169	BE079334	Hs.271630	ESTs	22.20
45	124281	AI333756	Hs.111801	arsenate resistance protein ARS2	42.20
	124472	N52517	Hs.102670	EST	32.60
	124617	AW628168	Hs.152684	ESTs	21.80
	124631	NM_014053	Hs.270694	FLVCR protein	30.40
	124839	R55784	Hs.140942	ESTs	21.20
50	125186	AA610620	Hs.181244	major histocompatibility complex, class	42.80
	125321	T86652	Hs.178294	ESTs	27.00
	125535	NM_013243	Hs.22215	secretogranin III	23.80
	125646	AA628962	Hs.75209	protein kinase (cAMP-dependent, catalyti	23.20
	125684	AW589427	Hs.158849	Homo sapiens cDNA: FLJ21663 fis, clone C	21.20
55	125724	AL360190	Hs.295978	Homo sapiens mRNA full length insert cDN	48.80
	125847	AW161885	Hs.249034	ESTs	31.00
	125934	AA193325	Hs.32646	hypothetical protein FLJ21901	21.20
	126077	M78772	Hs.210836	ESTs	49.80
	126299	AW979155	Hs.298275	amino acid transporter 2	21.80
60	126395	AI468004	Hs.278956	hypothetical protein FLJ12929	71.00
	126433	AA325606		gb:EST28707 Cerebellum II Homo sapiens c	23.20
	126509	R47400	Hs.23850	ESTs	23.80
	126538	AB030656	Hs.17377	coronin, actin-binding protein, 1C	23.10
	126666	AA648886	Hs.151999	ESTs	36.00
65	126812	AB037860	Hs.173933	nuclear factor I/A	20.80
	126872	AW450979		gb:UI-H-BI3-ata-a-12-0-UI.s1 NCI_CGAP_Su	46.29
	127046	AA321948	Hs.293968	ESTs	22.80
	127431	AW771958	Hs.175437	ESTs, Moderately similar to PC4259 ferri	30.00
	127489	AA650250	Hs.272076	ESTs	20.80
70	127521	AW297206	Hs.164018	ESTs	25.20
	127742	AW293496	Hs.180138	ESTs	28.00
	127925	AA805151	Hs.3628	mitogen-activated protein kinase kinase	21.20
	127930	AA809672	Hs.123304	ESTs	20.54
	127968	AA830201	Hs.124347	ESTs	28.20
75	127987	AI022103	Hs.124511	ESTs	19.60
	128116	H07103	Hs.286014	Homo sapiens, clone IMAGE:3867243, mRNA	20.40
	128609	NM_003616	Hs.102456	survival of motor neuron protein Interac	34.40
	128777	AI878918	Hs.10526	cysteine and glycine-rich protein 2	53.80
	128949	AA009647	Hs.8850	a disintegrin and metalloproteinase doma	23.00
80	129168	AI132988	Hs.109052	chromosome 14 open reading frame 2	37.60
	129404	AI267700	Hs.317584	ESTs	28.60
	129527	AA769221	Hs.270847	delta-tubulin	40.80
	129574	AA026815	Hs.11463	UMP-CMP kinase	31.20
	129598	N30436	Hs.11556	Homo sapiens cDNA FLJ12566 fis, clone NT	29.60
85	129785	H19006	Hs.184780	ESTs	72.20
	129970	AV655808	Hs.296198	chromosome 12 open reading frame 4	22.20

	130149	AW067805	Hs.172665	methylenetetrahydrofolate dehydrogenase	29.60
	130199	Z48579	Hs.172028	a disintegrin and metalloproteinase doma	27.60
	130441	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	28.36
5	130466	W19744	Hs.180059	Homo sapiens cDNA FLJ20653 fis, clone KA	20.20
	130482	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	22.40
	130617	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	19.60
	130703	R77776	Hs.18103	ESTs	19.40
	130732	AW690487	Hs.63984	cadherin 13, H-cadherin (heart)	21.40
10	130867	NM_001072	Hs.284239	UDP glycosyltransferase 1 family, polype	110.00
	131028	AJ879165	Hs.2227	CCAAT/enhancer binding protein (C/EBP),	25.20
	131086	AL035461	Hs.2281	chromogranin B (secretogranin 1)	40.60
	131284	NM_001429	Hs.25272	E1A binding protein p300	24.60
	131775	AB014548	Hs.31921	KIAA0648 protein	21.00
	131860	BE383676	Hs.334	Rho guanine nucleotide exchange factor (33.40
15	131945	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	60.80
	132040	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	20.40
	132084	NM_002267	Hs.3886	karyopherin alpha 3 (importin alpha 4)	29.40
	132389	AA310393	Hs.190044	ESTs	32.40
	132437	AA152106	Hs.4859	cyclin L ania-6a	27.40
20	132550	AW969253	Hs.170195	bone morphogenetic protein 7 (osteogenic	75.60
	132617	AF037335	Hs.5338	carbonic anhydrase XII	31.36
	132632	AU076916	Hs.5398	guanine monophosphate synthetase	32.40
	132672	W27721	Hs.54697	Cdc42 guanine exchange factor (GEF) 9	23.40
	132742	AA025480	Hs.292812	ESTs, Weakly similar to T33468 hypothei	61.20
25	132771	Y10275	Hs.56407	phosphoserine phosphatase	22.33
	133070	U92649	Hs.64311	a disintegrin and metalloproteinase doma	23.50
	133153	AF070592	Hs.66170	HSKM-B protein	30.00
	133181	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	23.80
	133282	AA449015	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye	51.60
30	133350	AI499220	Hs.71573	hypothetical protein FLJ10074	33.00
	133592	AV652066	Hs.75113	general transcription factor IIIA	82.00
	133658	AA319146	Hs.75426	secretogranin II (chromogranin C)	
	133665	AB011155	Hs.170290	discs, large (Drosophila) homolog 5	69.33
	134032	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor	33.20
35	134125	NM_014781	Hs.50421	KIAA0203 gene product	31.60
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	30.60
	134321	BE538082	Hs.8172	ESTs, Moderately similar to A45010 X-fin	23.40
	134367	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	49.20
	134570	U66615	Hs.172280	SWI/SNF related, matrix associated, acti	20.20
40	134753	NM_006482	Hs.173135	dual-specificity tyrosine-(Y)-phosphoryl	20.80
	135002	AA448542	Hs.251677	G antigen 7B	37.60
	135029	H58818	Hs.187579	hydroxysteroid (17-beta) dehydrogenase	53.40
	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	31.60
45	135345	X53655	Hs.99171	neurotrophin 3	28.80

TABLE 4B shows the accession numbers for those primekeys lacking unigenelD's for Table 4A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
123619	371681_1	AA602964 AA609200
126433	127143_1	AA325506 AA099517 N89423
126872	142696_1	AW450979 AA136653 AA136656 AW419381 AA984358 AA492073 BE168945 AA809054 AW238038 BE011212 BE011359
		BE011367 BE011368 BE011362 BE011215 BE011365 BE011363
106851	322947_1	AI458623 AA639708 AA485409 R22065 AA485570
118720	genbank_N73515	N73515
120515	genbank_AA258356	AA258356
117099	321871_1	H93599 H97976 H80036
101447	entrez_M21305	M21305
123130	genbank_AA487200	AA487200

Table 5A shows 680 genes up-regulated in squamous cell carcinoma or adenocarcinoma lung tumors relative to normal lung and chronically diseased lung. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number							
	ExAccn:	Exemplar Accession number, Genbank accession number							
	UnigeneID:	Unigene number							
	Unigene Title:	Unigene gene title							
10	R1:	70th percentile of AI for squamous cell carcinoma and adenocarcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R2:	80th percentile of AI adenocarcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R3:	80th percentile of AI squamous cell carcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R4:	80th percentile of AI adenocarcinoma lung tumor samples divided by the 80th percentile of AI for squamous cell carcinoma lung tumor samples.							
15	R5:	70th percentile of AI for squamous cell carcinoma and adenocarcinoma lung tumor samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by 90th percentile of AI for normal and chronically diseased lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples							
20	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2	R3	R4	R5
	100035			AFFX control: GAPDH					6.76
	100036			AFFX control: GAPDH					5.77
	100037			AFFX control: GAPDH					5.75
25	100071	A28102		Human GABAA receptor alpha-3 subunit		8.00			
	100114	X02308	Hs.82952	thymidylate synthetase					5.71
	100154	H60720	Hs.81892	KIAA0101 gene product	3.84				
	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3	3.33				
	100188	AW247090	Hs.57101	minichromosome maintenance deficient (S.					4.52
	100202	BE294407	Hs.99910	phosphofructokinase, platelet					5.49
30	100216	AA489908	Hs.1390	prolaserase (prosome, macropain) subunit,					5.67
	100269	NM_001949	Hs.1189	E2F transcription factor 3	2.55				
	100287	AU076657	Hs.1600	chaperonin containing TCP1, subunit 5 (e					5.66
	100297	AU077258	Hs.182429	protein disulfide isomerase-related prot					3.81
	100330	AW410976	Hs.77152	minichromosome maintenance deficient (S.					4.50
35	100335	AW247529	Hs.6793	platelet-activating factor acetylhydrola	5.07				
	100360	W70171	Hs.75939	uridine monophosphate kinase					4.82
	100372	NM_014791	Hs.184339	KIAA0175 gene product					3.79
	100474	NM_000699	Hs.300280	amylase, alpha 2A; pancreatic				15.65	
40	100486	T19006	Hs.10842	RAN, member RAS oncogene family					5.49
	100491	D56165	Hs.275163	non-metastatic cells 2, protein (NM23B)					4.17
	100516	D90278	Hs.11	carcinoembryonic antigen-related cell ad		7.20			
	100522	X51501	Hs.99949	prolactin-induced protein				14.20	
	100559	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys	3.10				
	100576	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid				9.30	
45	100629	AA015693	Hs.21291	mitogen-activated protein kinase kinase				20.60	
	100661	BE623001	Hs.132748	Homo sapiens ribosomal protein L39 mRNA,	3.85				
	100677	AA353686	Hs.57813	zinc ribbon domain containing, 1		8.60			
	100696	D14887	Hs.121685	general transcription factor IIA, 1 (37k				10.00	
	100709	N26539	Hs.100469	myeloid/lymphoid or mixed-lineage leukem			24.80		
50	100761	BE208491	Hs.295112	KIAA0618 gene product		7.60			
	100830	AC004770	Hs.47556	flap structure-specific endonuclease 1					7.99
	100867	U14822		gb:Human transketolase-like protein gene		10.20			
	100902	M16029	Hs.287270	ret proto-oncogene (multiple endocrine n		8.00			
	100906	AU076916	Hs.5398	guanine monophosphate synthetase					5.16
55	100960	J00124	Hs.117729	keratin 14 (epidermolysis bullosa simple	2.57				
	101045	J05614		gb:Human proliferating cell nuclear anti					4.69
	101061	NM_000175	Hs.180532	glucose phosphate isomerase					4.19
	101071	L02840	Hs.84244	potassium voltage-gated channel, Shab-re		12.91			
	101124	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	3.12				
60	101175	U82671	Hs.36980	melanoma antigen, family A, 2	3.50				
	101181	BE262621	Hs.73798	macrophage migration inhibitory factor (5.69
	101204	L24203	Hs.82237	ataxia-telangiectasia group D-associated	4.08				
	101210	L29301	Hs.2353	oploid receptor, mu 1			6.40		
	101216	AA284166	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	2.53				
65	101228	AA333387	Hs.82916	chaperonin containing TCP1, subunit 6A (7.90
	101233	AL135173	Hs.878	sorbitol dehydrogenase					4.45
	101273	Z11933	Hs.182505	POU domain, class 3, transcription facto	8.50				
	101342	U52112	Hs.182018	interleukin-1 receptor-associated kinase					4.17
	101346	AJ738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N				21.89	
70	101369	NM_000892	Hs.1901	kallikrein B, plasma (Fletcher factor) 1				12.80	
	101396	BE267931	Hs.78996	proliferating cell nuclear antigen	3.24				
	101431	BE185289	Hs.1076	small proline-rich protein 1B (oomifin)					7.90
	101448	NM_000424	Hs.195850	keratin 5 (epidermolysis bullosa simplex	8.31				
	101462	AL035668	Hs.73853	bone morphogenetic protein 2				38.80	
75	101466	BE262660	Hs.170197	glutamic-oxaloacetic transaminase 2, mit					4.01
	101484	AA053486	Hs.20315	interferon-induced protein with letratri				12.00	
	101502	M26958		gb:Human parathyroid hormone-related pro	10.50				
	101505	AA307680	Hs.75692	asparagine synthetase					4.46
	101526	NM_002197	Hs.154721	aconitase 1, soluble	4.02				
80	101535	X57152	Hs.99853	fibrillarin					4.65
	101577	M34353	Hs.1041	v-ros avian UR2 sarcoma virus oncogene h				9.09	
	101649	AW959908	Hs.1690	heparin-binding growth factor binding pr	54.00				
	101663	NM_003528	Hs.2178	H2B histone family, member Q	5.59				
	101664	AA436989	Hs.121017	H2A histone family, member A	7.00				
85	101669	L24498	Hs.80409	growth arrest and DNA-damage-inducible,		7.60			

	101695	M69136	Hs.135626	chymase 1, masl cell	4.79		
	101724	L11690	Hs.620	bulbosus pemphigoid antigen 1 (230/240kD)	15.21		
	101748	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen)	55.50		
5	101759	M80244	Hs.184601	solute carrier family 7 (cationic amino)		18.57	4.10
	101771	NM_002432	Hs.153837	myeloid cell nuclear differentiation ant			
	101804	M86699	Hs.169840	TTK protein kinase	4.50		
	101809	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	140.00		
	101833	AU076442	Hs.117938	collagen, type XVII, alpha 1	2.56		
10	101842	M93221	Hs.75182	mannose receptor, C type 1		12.80	
	101851	BE260964	Hs.82045	midkine (neurite growth-promoting factor			5.88
	102002	NM_002484	Hs.81469	nucleotide binding protein 1 (E.coli Min	7.80		
	102039	AL134223	Hs.306098	aldo-keto reductase family 1, member C1			4.35
	102072	U09410	Hs.78743	zinc finger protein 131 (clone pHZ-10)	7.40		
	102083	T35901	Hs.75117	interleukin enhancer binding factor 2, 4			5.12
15	102111	L36196	Hs.81884	sulfotransferase family, cytosolic, 2A,		12.00	
	102123	NM_001809	Hs.1594	centromere protein A (17kD)	6.20		
	102154	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kallinin	2.62		
	102193	AL036535	Hs.313	secreted phosphoprotein 1 (osteopontin,	5.85		
20	102217	AA829978	Hs.301613	JTV1 gene			6.18
	102224	NM_002810	Hs.148495	proteasome (prosome, macropain) 26S subu			4.49
	102234	AW163390	Hs.278554	heterochromatin-like protein 1			5.80
	102251	NM_004398	Hs.41706	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	4.50		
	102305	AL043202	Hs.90073	chromosome segregation 1 (yeast homolog)			5.15
25	102330	BE298063	Hs.77254	chromobox homolog 1 (Drosophila HP1 beta			4.17
	102340	U37055	Hs.278557	macrophage stimulating 1 (hepatocyte gro		9.33	
	102348	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member	8.87		
	102358	U39817	Hs.36820	Bloom syndrome	15.91		
30	102394	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma		19.20	
	102404	NM_005429	Hs.79141	vascular endothelial growth factor C			14.00
	102537	U57094	Hs.50477	RAB27A, member RAS oncogene family			12.00
	102581	AU077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2			4.57
	102605	AJ435128	Hs.181369	ubiquitin fusion degradation 1-like			3.98
	102610	U65011	Hs.30743	preferentially expressed antigen in meta	77.50		
35	102623	AW249285	Hs.37110	melanoma antigen, family A, 9	12.50		
	102642	AA205847	Hs.23016	G protein-coupled receptor		22.00	
	102654	AV649989	Hs.24385	Human hbc547 mRNA sequence	12.00		
	102659	BE245169	Hs.211610	CUG triplet repeat, RNA-binding protein			12.80
	102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	6.50		
40	102672	U72066	Hs.29287	retinoblastoma-binding protein 8	8.50		
	102687	NM_007019	Hs.93002	ubiquitin carrier protein E2-C			9.24
	102696	BE540274	Hs.239	forkhead box M1			5.54
	102768	U82321		gb:Homo sapiens clone 14.9B mRNA sequenc	6.60		
45	102781	BE258778	Hs.108809	chaperonin containing TCP1, subunit 7 (e			3.78
	102784	U85658	Hs.61795	transcription factor AP-2 gamma (activat		14.40	4.26
	102824	U90916	Hs.82845	Homo sapiens cDNA: FLJ21930 tis, clone H			
	102829	NM_006183	Hs.80962	neurotensin	8.00		
	102888	AJ346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1			5.50
50	102892	BE440042	Hs.83326	matrix metalloproteinase 3 (stromelysin		6.70	
	102913	NM_002275	Hs.80342	keratin 15	4.64		
	102935	BE561850	Hs.80506	small nuclear ribonucleoprotein polypept	2.93		
	102951	X15218	Hs.2969	v-ski avian sarcoma viral oncogene homol		11.40	
	102983	BE387202	Hs.118638	non-melanocytic cells 1, protein (NM23A)			7.26
55	103023	AW500470	Hs.117950	multifunctional polypeptide similar to S	3.01		
	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	27.90		
	103038	AA926960	Hs.334883	CDC28 protein kinase 1			8.79
	103060	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin			4.27
	103099	AJ693251	Hs.8248	NADH dehydrogenase (ubiquinone) Fo-S pro	9.80		
60	103119	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	4.05		
	103168	X53463	Hs.2704	glutathione peroxidase 2 (gastrointestin	3.07		
	103185	NM_006825	Hs.74368	transmembrane protein (63kD), endoplasmic			5.62
	103192	M22440	Hs.170009	transforming growth factor, alpha	7.40		
	103223	BE275607	Hs.1708	chaperonin containing TCP1, subunit 3 (g			4.70
	103242	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o		100.00	
65	103316	X83301	Hs.324728	SMA5			9.80
	103375	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	9.71		
	103376	AL036166	Hs.323378	coated vesicle membrane protein	14.00		
	103385	NM_007069	Hs.37189	similar to rat HREV107			11.00
	103391	X94453	Hs.114366	pyrroline-5-carboxylate synthetase (glut	2.93		
70	103404	BE394784	Hs.78596	proteasome (prosome, macropain) subunit,			5.15
	103430	BE564090	Hs.20716	translocase of inner mitochondrial membr			3.98
	103446	X98834	Hs.79971	sal (Drosophila)-like 2		21.40	
	103476	Y07701	Hs.293007	aminopeptidase puromycin sensitive	13.00		
	103477	AJ011812	Hs.119018	transcription factor NRF		6.40	
75	103478	BE514982	Hs.38991	S100 calcium-binding protein A2	5.02		
	103515	Y10275	Hs.56407	phosphoserine phosphatase	10.50		
	103558	BE816547	Hs.2785	keratin 17	6.41		
	103580	AA328046	Hs.46405	polymerase (RNA) II (DNA directed) polyp			3.84
	103587	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	78.50		
80	103594	AJ368580	Hs.816	SRY (sex determining region Y)-box 2	6.51		
	103636	NM_006235	Hs.2407	POU domain, class 2, associating factor	3.50		
	103768	AF065009		gb:Homo sapiens full length insert cDNA			4.48
	103841	AA314821	Hs.38178	hypothetical protein FLJ23468	8.00		
	103847	AF219946	Hs.102237	tubby super-family protein	10.40		
85	103913	AW967500	Hs.133543	ESTs		15.60	
	104094	AA418187	Hs.330515	ESTs	6.60		

WO 02/086443			PCT/US02/12476		
5	104150	AL122044	Hs.331633	hypothetical protein DKFZp566N034	26.00
	104257	BE560621	Hs.9222	estrogen receptor binding site associate	6.80
	104261	AW248364	Hs.5409	RNA polymerase I subunit	3.98
	104331	AB040450	Hs.279862	cdk inhibitor p21 binding protein	6.80
	104415	BE410992	Hs.258730	heme-regulated initiation factor 2-alpha	10.29
	104558	RS5678	Hs.88959	hypothetical protein MGC4816	4.21
	104590	AW373062	Hs.83623	nuclear receptor subfamily 1, group I, m	15.79
	104658	AA360954	Hs.27268	Homo sapiens cDNA: FLJ21933 fis, clone H	17.40
10	104660	BE298665	Hs.14846	Homo sapiens mRNA: cDNA DKFZp564D016 (fr	6.40
	104689	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	6.55
	104754	AI206234	Hs.155924	cAMP responsive element modulator	10.00
	104758	BE560269	Hs.7010	NP0002 protein	4.47
	104971	BE311926	Hs.15830	hypothetical protein FLJ12691	2.87
	105011	BE091926	Hs.16244	mitotic spindle coiled-coil related prot	3.83
15	105012	AF098158	Hs.9329	chromosome 20 open reading frame 1	2.86
	105026	AA809485	Hs.124219	hypothetical protein FLJ12934	11.00
	105076	AI598252	Hs.37810	hypothetical protein MGC14833	5.01
	105132	AA148164	Hs.247280	HBV associated factor	3.99
	105143	AI368836	Hs.24808	ESTs, Weakly similar to I38022 hypotheti	11.00
20	105158	AW976357	Hs.234545	hypothetical protein NUF2R	16.00
	105175	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	4.32
	105200	AA328102	Hs.24641	cytoskeleton associated protein 2	3.00
	105264	AA227934		gbzr57e08.s1 Soares_NhHMPu_S1 Homo sapi	10.00
25	105298	BE387790	Hs.26369	hypothetical protein FLJ20287	3.69
	105409	AW505076	Hs.301855	DiGeorge syndrome critical region gene 8	9.20
	105460	AW296078	Hs.271721	Homo sapiens, clone IMAGE:4179986, mRNA,	7.80
	105667	AA767526	Hs.22030	paired box gene 5 (B-cell lineage specif	4.12
	105743	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	3.82
	105782	H09748	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro	27.00
30	105848	AW954064	Hs.24951	ESTs	7.60
	105891	U55984	Hs.289088	heat shock 90kD protein 1, alpha	4.14
	106019	AF221993	Hs.46743	McKusick-Kaufman syndrome	16.80
	106069	BE566623	Hs.29899	ESTs, Weakly similar to G02075 transcrip	23.40
	106073	AL157441	Hs.17834	downstream neighbor of SON	9.50
35	106126	AA576953	Hs.22972	hypothetical protein FLJ13352	6.00
	106159	AK001301	Hs.3487	hypothetical protein FLJ10439	
	106220	D61329	Hs.32196	mitochondrial ribosomal protein L36	13.20
	106260	AI097144	Hs.5250	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.02
40	106300	Y10043	Hs.19114	high-mobility group (nonhistone chromoso	6.60
	106307	AA436174	Hs.37751	ESTs, Weakly similar to putative p150 (
	106318	AA025610	Hs.9605	cleavage and polyadenylation specific fa	5.04
	106341	AF191020	Hs.5243	hypothetical protein, estradiol-induced	7.25
	106440	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub	13.80
	106481	D61594	Hs.17279	tyrosylprotein sulfotransferase 1	4.75
45	106586	AA243837	Hs.57787	ESTs	10.84
	106605	AW772299	Hs.21103	Homo sapiens mRNA: cDNA DKFZp564B076 (fr	45.60
	106654	AW075485	Hs.286049	phosphoserine aminotransferase	28.00
	106785	Y15227	Hs.20149	deleted in lymphocytic leukemia, 1	3.00
	106813	C05766	Hs.181022	CGI-07 protein	11.40
50	106895	AK001826	Hs.25245	hypothetical protein FLJ11269	6.00
	106913	AI219346	Hs.86178	M-phase phosphoprotein 9	6.56
	106919	AW043637	Hs.21766	ESTs, Weakly similar to ALU5_HUMAN ALU S	4.27
	107054	AI076459	Hs.15978	KIAA1272 protein	34.80
55	107059	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Ra	4.71
	107098	AI823593	Hs.27688	ESTs	24.80
	107104	AU076640	Hs.15243	nucleolar protein 1 (120kD)	7.05
	107129	AC004770	Hs.4756	flap structure-specific endonuclease 1	2.60
	107198	AV657225	Hs.9846	KIAA1040 protein	19.20
60	107203	D20426	Hs.41639	programmed cell death 2	7.60
	107217	AL080235	Hs.35861	DKFZP586E1621 protein	9.50
	107284	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte	2.71
	107318	T74445	Hs.5957	Homo sapiens clone 24416 mRNA sequence	8.71
	107516	X57152	Hs.99853	fibrillarin	4.33
65	107529	BE515065	Hs.296585	nucleolar protein (KKE/D repeat)	4.00
	107728	AA019551	Hs.294151	Homo sapiens, clone IMAGE:3603836, mRNA,	10.80
	107851	AA022953	Hs.61172	EST	8.00
	107901	L42612	Hs.335952	keratin 6B	3.40
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	2.88
70	107932	AW392555	Hs.18878	hypothetical protein FLJ21620	7.50
	108015	AW298357	Hs.49927	protein kinase NYD-SP15	23.40
	108056	AA043675	Hs.62633	ESTs	12.80
	108075	AI867370	Hs.139709	hypothetical protein FLJ12572	12.80
	108187	BE245374	Hs.27842	hypothetical protein FLJ11210	7.00
	108296	N31256	Hs.161623	ESTs	6.60
75	108305	AA071391		gbzm61e06.r1 Stratagene fibroblast (937	11.80
	108393	AA075211		gbzm86a08.r1 Stratagene ovarian cancer	11.80
	108480	AL133092	Hs.68055	hypothetical protein DKFZp4340428	20.80
	108554	AA084948		gbzm13b09.s1 Stratagene hNT neuron (937	6.40
	108573	AA086005		gbzm84c04.s1 Stratagene colon (937204)	25.40
80	108584	AA088326	Hs.120905	Homo sapiens cDNA FLJ11448 fis, clone HE	9.60
	108597	AK000292	Hs.278732	hypothetical protein FLJ20285	14.60
	108695	AB029000	Hs.70823	KIAA1077 protein	3.00
	108699	AA121514	Hs.70832	ESTs	10.00
	108700	AA121518	Hs.193540	ESTs, Moderately similar to 2109260A B c	11.00
85	108780	AU076442	Hs.117938	collagen, type XVII, alpha 1	11.21

	108810	AW295647	Hs.71331	hypothetical protein MGC5350	8.50			
	108816	AA130884	Hs.270501	ESTs, Moderately similar to ALU2_HUMAN		7.40		
	108857	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	4.00			
	108860	AA133334	Hs.129911	ESTs	6.09			
5	108937	AL050107	Hs.24341	transcriptional co-activator with PDZ-bl	3.00			
	109010	NM_007240	Hs.44229	dual specificity phosphatase 12	2.69			
	109121	BE389387	Hs.49767	NADH dehydrogenase (ubiquinone) Fe-S pro			4.53	
	109166	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rakines)	10.58			
	109227	AA766998	Hs.85874	Human DNA sequence from clone RP11-16L21		9.00		
10	109415	U80736	Hs.110826	trinucleotide repeat containing 9		51.40		
	109418	AI866946	Hs.161707	ESTs			11.00	
	109454	AA232255	Hs.295232	ESTs, Moderately similar to A46010 X-li		17.60		
	109502	AW967069	Hs.211556	hypothetical protein MGC5487		9.49		
	109543	AA584994	Hs.222851	ESTs		12.67		
15	109548	H17800	Hs.7154	ESTs			10.40	
	109680	AB037734	Hs.4993	KIAA1313 protein		33.20		
	109700	F09609		gb:HSC33H092 normalized infant brain cDN			16.00	
	109704	AI743880	Hs.12876	ESTs		11.00		
	109792	R49625		gb:yg61f03.s1 Soares infant brain 1N1B H			12.60	
20	109981	BE546208	Hs.26090	hypothetical protein FLJ20272	4.00			
	109998	AL042201	Hs.21273	transcription factor NYD-sp10		7.80		
	110039	H11938	Hs.21907	histone acetyltransferase		7.00		
	110156	AA581322	Hs.4213	hypothetical protein MGC16207			4.24	
	110500	AA907723	Hs.36952	ESTs	4.50			
25	110551	AW450381	Hs.14529	ESTs		8.60		
	110561	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	3.06			
	110854	BE612992	Hs.27931	hypothetical protein FLJ10607 similar to		6.80		
	110886	AW274992	Hs.72249	three-PDZ containing protein similar to		8.80		
	110916	BE178102	Hs.24349	ESTs		6.80		
30	111003	N52980	Hs.83765	dihydrofolate reductase			16.80	
	111337	AA837396	Hs.263925	LIS1-interacting protein NUDE1, rat homo	2.54			
	111434	R01608	Hs.142738	ESTs			9.80	
	111439	AI476429	Hs.19238	ESTs			10.40	
	111540	U82670	Hs.9786	zinc finger protein 275		15.40		
35	111597	R11499	Hs.189716	ESTs			9.20	
	111895	T80581	Hs.12723	Homo sapiens clone 25153 mRNA sequence	6.80			
	111929	AF027208	Hs.112360	prominin (mouse)-like 1			14.67	
	112054	R43590		gb:yc85g02.s1 Soares infant brain 1N1B H	10.80			
	112210	R49645	Hs.7004	ESTs			10.20	
40	112244	AB029000	Hs.70823	KIAA1077 protein	2.99			
	112382	R59904		gb:yh07g12.s1 Soares infant brain 1N1B H		6.60		
	112392	R60763	Hs.193274	ESTs, Moderately similar to I57588 HSrel		7.10		
	112442	AA280174	Hs.285681	Williams-Beuren syndrome chromosome reg)	3.00			
	112539	R70318	Hs.339730	ESTs			37.20	
45	112772	AI992283	Hs.35437	ESTs, Moderately similar to I38026 MLN 6			14.60	
	112869	BE261750	Hs.4747	dyskeratosis congenita 1, dyskerin				4.83
	112935	R71449	Hs.268760	ESTs	2.73			
	112970	AA694010	Hs.6932	Homo sapiens clone 23809 mRNA sequence			12.00	
50	112973	AB033023	Hs.318127	hypothetical protein FLJ10201	11.50			
	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f		10.89		
	113063	W15573	Hs.5027	ESTs, Weakly similar to A47582 B-cell gr	15.00			
	113073	N39342	Hs.103042	mitochondria-associated protein 1B		15.31		
	113078	T40444	Hs.118354	CAT56 protein	7.00			
	113238	R45467	Hs.189813	ESTs			41.20	
55	113591	T91881	Hs.200597	KIAA0563 gene product			9.40	
	113702	T97307		gb:ye53h05.s1 Soares fetal liver spleen	25.00			
	113844	AI369275	Hs.243010	Homo sapiens cDNA FLJ14445 fis, clone HE			13.91	
	113984	R96696	Hs.35598	ESTs		7.80		
	114073	R44953	Hs.22908	Homo sapiens mRNA; cDNA DKFZp434J1027 (f		7.20		
60	114162	AF155661	Hs.22265	pyruvate dehydrogenase phosphatase	3.42			
	114208	AL049466	Hs.7859	ESTs		6.74		
	114251	H15261	Hs.21948	ESTs			33.20	
	114285	R44338	Hs.22974	ESTs			13.20	
	114313	H18456	Hs.27946	ESTs			10.00	
65	114339	AA782845	Hs.22790	ESTs		7.80		
	114407	BE539976	Hs.103305	Homo sapiens mRNA; cDNA DKFZp434B0425 (f				4.14
	114560	AI452469	Hs.165221	ESTs			9.80	
	114699	AA127388		gb:zn90d09.r1 Stratagene lung carcinoma		7.60		
70	114767	AI859865	Hs.154443	minichromosome maintenance deficient (S	3.21			
	114793	AA158245		gb:zo76c03.s1 Stratagene pancreas (93720		6.00		
	114833	AI417215	Hs.87159	hypothetical protein FLJ12577			11.40	
	115047	BE270930	Hs.82916	chaperonin containing TCP1, subunit 6A (4.31
	115060	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3				4.03
	115097	AA256213	Hs.72010	ESTs			35.40	
75	115113	AA256460		gb:zr81a04.s1 Soares_NhHMPu_S1 Homo sapi			15.20	
	115123	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m				4.19
	115134	AW968073	Hs.194331	ESTs, Highly similar to A55713 inositol			12.40	
	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	25.00			
	115347	AA356792	Hs.334824	hypothetical protein FLJ14825		7.00		
80	115414	AA662240	Hs.283099	AF15q14 protein	3.25			
	115522	BE614387	Hs.333893	c-Myc target JPO1	3.68			
	115536	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	10.50			
	115566	AI142336	Hs.43977	Human DNA sequence from clone RP11-195N1			24.40	
	115645	AI207410	Hs.69280	Homo sapiens, clone IMAGE:3636299, mRNA,	4.17			
85	115648	AW016811	Hs.234478	Homo sapiens cDNA: FLJ22648 fis, clone H		6.00		

5	115652	BE093589	Hs.38178	hypothetical protein FLJ23468	3.81			
	115697	D31382	Hs.63325	transmembrane protease, serine 4	62.14			
	115793	AA424883	Hs.70333	hypothetical protein MGC10753			11.80	
	115816	BE042915	Hs.287588	Homo sapiens cDNA FLJ13675 fis, clone PL			9.71	
	115892	AA291377	Hs.50831	ESTs		27.40		
10	115906	AI767756	Hs.82302	Homo sapiens cDNA FLJ14814 fis, clone NT	2.53			
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	11.82			
	115965	AA001732	Hs.173233	hypothetical protein FLJ10970			34.29	
	115978	AI035864	Hs.69517	cDNA for differentially expressed CO16 g			8.23	
	115985	AA447709	Hs.268115	ESTs, Weakly similar to T08599 probable	3.00			
15	116090	AI591147	Hs.61232	ESTs	5.17			
	116096	AA582382	Hs.59982	ESTs		8.20		
	116127	AF126743	Hs.279884	DNAJ domain-containing		10.60		
	116157	BE438338	Hs.44298	mitochondrial ribosomal protein S17.			5.82	
	116190	AI949095	Hs.67776	ESTs, Weakly similar to T22341 hypotheti			4.08	
20	116278	NM_003686	Hs.47504	exonuclease 1	9.50			
	116335	AK001100	Hs.41690	desmoccilin 3	3.67			
	116496	AW450694	Hs.21433	hypothetical protein DKFZp547J036		7.00		
	116503	AI925316	Hs.212617	ESTs			12.60	
	116674	AI768015	Hs.92127	ESTs		32.00		
25	116929	AA586922	Hs.80475	polymerase (RNA) II (DNA directed) polyp		7.60		
	116973	AI702054	Hs.166982	phosphatidylinositol glycan, class F	9.80			
	116993	AA17023	Hs.40478	ESTs			10.20	
	117079	H92325		gbys85R05.s1 Soares retina N2b4HR Homo			15.20	
	117317	AI253517	Hs.43322	ESTs			13.40	
30	117326	N23629	Hs.241420	Homo sapiens mRNA for KIAA1756 protein,			20.60	
	117396	W20128	Hs.296039	ESTs			10.60	
	117412	N32536	Hs.42645	ESTs			16.00	
	117519	N32528	Hs.146286	kinesin family member 13A			9.11	
	117693	AW179019	Hs.112110	mitochondrial ribosomal protein L42			4.01	
35	117721	N45100	Hs.93939	EST			19.80	
	117881	AF161470	Hs.260622	butyrate-induced transcript 1	2.71			
	117903	AA768283	Hs.47111	ESTs			17.80	
	117992	AI015709	Hs.172089	Homo sapiens mRNA; cDNA DKFZp586I2022 (f			4.17	
	118013	AI674126	Hs.94031	ESTs			10.60	
40	118017	AI813444	Hs.42197	ESTs		8.82		
	118186	N22886	Hs.42380	ESTs	7.00			
	118325	AI868065	Hs.166184	intersectin 2			13.80	
	118367	N64269	Hs.48946	EST		6.14		
	118368	N64339	Hs.48956	gap junction protein, beta 6 (connexin 3	3.14			
45	118472	AI157545	Hs.42179	bromodomain and PHD finger containing, 3		12.40		
	118709	AA232970	Hs.293774	ESTs			12.20	
	119025	BE003760	Hs.55209	Homo sapiens mRNA; cDNA DKFZp434K0514 (f	4.50			
	119027	AF086161	Hs.114611	hypothetical protein FLJ11808	3.22			
	119052	R10889		gbcyf38d02.s1 Soares fetal liver spleen	9.60			
50	119164	AF221993	Hs.46743	McKusick-Kaufman syndrome		6.60		
	119186	AI979147	Hs.101265	hypothetical protein FLJ22593			10.80	
	119243	T12603		gb:CHR90123 Chromosome 9 exon II Homo sa			9.44	
	119490	AA195276	Hs.263858	ESTs, Moderately similar to B34087 hypot			11.80	
	119499	AI918906	Hs.55080	ESTs		14.80		
55	119599	W45552		gbzcc26d03.s1 Soares_senescent_fibroblas		12.60		
	119780	NM_016625	Hs.191381	hypothetical protein	17.00			
	119845	W79123	Hs.58561	G protein-coupled receptor 87	13.50			
	119941	AA699485	Hs.58896	ESTs		8.00		
	119994	AA642402	Hs.59142	ESTs	7.73			
60	120102	W67353	Hs.170218	KIAA0251 protein		39.60		
	120104	AK000123	Hs.180479	hypothetical protein FLJ20116	2.91			
	120294	AK000059	Hs.153881	Homo sapiens NY-REN-62 antigen mRNA, par		8.20		
	120486	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	8.73			
	120599	AA804448	Hs.104463	ESTs	7.00			
65	120699	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos			10.00	
	120715	AA292700		gbzcs59a06.s1 NCL_CGAP_GCB1 Homo sapiens	9.40			
	120821	Y19062	Hs.96870	stauferin (Drosophila, RNA-binding protein			13.80	
	120859	AA826434	Hs.1619	achaela-scute complex (Drosophila) homol		9.00		
	120880	AA360240	Hs.97019	EST	15.60			
70	120983	AA389809	Hs.97587	EST		27.66		
	121034	AL389951	Hs.271623	nucleoporin 50kD		20.80		
	121121	AA399371	Hs.189095	similar to SALL1 (sal (Drosophila)-like	22.80			
	121313	AA402713	Hs.97872	ESTs			10.00	
	121369	AW450737	Hs.128791	CGI-09 protein	25.71			
75	121376	AA448103	Hs.187958	solute carrier family 6 (neurotransmitte			5.42	
	121476	AA412311	Hs.97903	ESTs	8.30			
	121509	AA868939	Hs.97888	ESTs	8.59			
	121553	AA412488	Hs.48820	TATA box binding protein (TBP)-associat	18.50			
	121753	AK000652	Hs.323518	WD repeat domain 5	7.00			
80	121838	AA425680	Hs.98441	ESTs			10.40	
	121857	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	6.00			
	121991	AA430058	Hs.98649	EST			12.20	
	122089	AW016543	Hs.98682	hypothetical protein FKSG32		8.60		
	122105	AW241685	Hs.98699	ESTs	6.14			
85	122163	AA435702	Hs.98829	EST			10.40	
	122318	AA429743		gbzv60b05.r1 Soares_testis_NHT Homo sap			18.20	
	122335	AA443258	Hs.241551	chloride channel, calcium activated, fam	13.50			
	122338	AA443311	Hs.98998	ESTs	4.80			
	122414	AI313473	Hs.99087	ESTs, Weakly similar to S47073 finger pr	8.00			

	122512	AF053305	Hs.98658	budding uninhibited by benzimidazoles 1	8.80		
	122516	AA449352	Hs.99217	ESTs		9.40	
	122702	AJ220089	Hs.99439	ESTs	9.20		
	122852	AJ580056	Hs.98992	ESTs		10.40	
5	122925	AW268962	Hs.111335	ESTs	6.80		
	123005	AW369771	Hs.52620	integrin, beta 8		12.60	
	123044	AK001035	Hs.130981	B-cell CLL/lymphoma 11A (zinc finger pro			5.35
	123160	AA489687	Hs.284235	ESTs, Weakly similar to I38022 hypotheti		6.06	
10	123315	AA496369		gbzv37d10.s1 Soares ovary tumor NbHOT H		12.40	
	123329	Z47542	Hs.179312	small nuclear RNA activating complex, po		11.80	
	123497	AA765256	Hs.135191	ESTs, Weakly similar to unnamed protein	12.00		
	123518	AL035414	Hs.21068	hypothetical protein		13.00	
	123519	AW015887	Hs.112574	ESTs	12.20		
	123514	AK000492	Hs.98806	hypothetical protein		7.80	
15	123616	AA680003	Hs.109363	Homo sapiens cDNA: FLJ23603 fis, clone L			10.60
	123673	BE550112	Hs.158549	ESTs, Weakly similar to T2D3_HUMAN TRANS	23.00		
	123727	AI083986	Hs.282977	hypothetical protein FLJ13490	7.00		
	123731	AA609839		gbcae52f01.s1 Stratagene lung carcinoma		9.80	
	123752	AA227714	Hs.179703	KIAA0129 gene product	3.50		
20	123900	AA621223	Hs.112953	EST			12.80
	124006	AI147155	Hs.270016	ESTs	97.00		
	124059	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti	3.02		
	124069	AF134160	Hs.7327	claudin 1		27.80	
	124191	T96509	Hs.248549	ESTs, Moderately similar to S65657 alpha			35.80
25	124273	AA457211	Hs.8858	bromodomain adjacent to zinc finger doma	7.20		
	124297	AL080215	Hs.102301	Homo sapiens mRNA; cDNA DKFZp586J0323 (f			11.00
	124305	AW963221		gb:EST375294 MAGE resequences, MAGH Homo			16.00
	124676	AJ360119.comp	Hs.181013	phosphoglycerate mutase 1 (brain)			6.08
	124874	BE550182	Hs.127826	RalGEF-like protein 3, mouse homolog			21.00
30	124904	AK000483	Hs.93872	KIAA1682 protein	9.40		
	124969	AI650360	Hs.100256	ESTs			10.80
	125000	T58615	Hs.110640	ESTs			9.80
	125201	AA693960	Hs.103158	ESTs, Weakly similar to T33296 hypotheti	7.60		
	125266	W90022	Hs.186809	ESTs, Highly similar to LCT2_HUMAN LEUKO	6.59		
35	125299	T32982	Hs.102720	ESTs			9.57
	125356	AJ057052	Hs.133554	ESTs, Weakly similar to Z195_HUMAN ZINC			14.00
	125370	AA256743	Hs.134158	Homo sapiens, Similar to KIAA0092 gene p	8.20		
	125418	AA777690	Hs.188501	ESTs			13.20
	125433	AL162066	Hs.54320	hypothetical protein DKFZp762D096	21.40		
40	125437	AI609449	Hs.140197	ESTs	6.98		
	125446	BE219987	Hs.166982	phosphatidylinositol glycan, class F	8.80		
	125711	AA305800	Hs.5672	hypothetical protein AF140225		11.20	
	125756	BE174587	Hs.289721	growth arrest specific transcript 5			4.31
	125757	AJ274906	Hs.166835	ESTs, Highly similar to 1814460A p53-ass			15.60
45	125769	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	3.20		
	125839	AW836261	Hs.337717	ESTs	8.20		
	125850	W85858	Hs.99804	ESTs	2.65		
	125875	H14480		gbym18b09.r1 Soares infant brain 1N1B H	7.40		
	125924	BE272506	Hs.82109	syndecan 1			4.23
50	125972	AJ927475	Hs.35406	ESTs, Highly similar to unnamed protein			3.98
	126034	H60340		gbv39b04.r1 Soares fetal liver spleen		10.60	
	126327	AA432266	Hs.44548	ESTs	11.60		
	126345	N49713		gbzyv23f06.s1 Soares fetal liver spleen	6.67		
	126435	AW614529	Hs.285947	CGI-19 protein		10.60	
55	126487	AA283909	Hs.184601	solute carrier family 7 (cationic amino			4.38
	126521	AJ475110	Hs.203933	ESTs	6.60		
	126522	W31912		gbzcv76d03.s1 Pancreatic Islet Homo sapi		14.80	
	126543	AL035864	Hs.69517	cDNA for differentially expressed CO16 g			4.01
	126567	AA058394	Hs.57887	ESTs, Weakly similar to KIAA0758 protein	7.80		
60	126605	AA676910		gbzj65h07.s1 Soares fetal liver spleen		11.60	
	126627	AA497044	Hs.20887	hypothetical protein FLJ10392		14.60	
	126628	N49776	Hs.170994	hypothetical protein MGC10946	8.00		
	126737	AW976516	Hs.283707	Homo sapiens cDNA: FLJ21354 fis, clone C	2.92		
	126795	AW975076	Hs.172569	nuclear phosphoprotein similar to S. cer	7.50		
65	126802	AW805510	Hs.97056	hypothetical protein FLJ21634	11.60		
	126892	AF121856	Hs.284291	sorting nexin 6	3.50		
	126928	AA480902	Hs.137401	ESTs			22.83
	126979	AA210954		gbzq89h10.r1 Stratagene hNT neuron (937			11.80
	126986	AJ279892	Hs.46801	sorting nexin 14			11.60
70	126992	AJ809521		gbwrf30e03.x1 Soares_NFL_T_GBC_S1 Homo s			20.80
	127066	R25066		gbzyg42c07.r1 Soares infant brain 1N1B H			27.60
	127099	AA347668		gb:EST54026 Fetal heart II Homo sapiens			21.60
	127139	AA830233	Hs.293585	ESTs			11.20
	127209	AA305023	Hs.81964	SEC24 (S. cerevisiae) related gene famil	3.10		
75	127221	BE062109	Hs.241551	chloride channel, calcium activated, fam	2.76		
	127225	AA315933	Hs.120879	ESTs			16.80
	127313	AK002014	Hs.47546	Homo sapiens cDNA FLJ11458 fis, clone HE	14.00		
	127444	AW978474	Hs.7560	Homo sapiens mRNA for KIAA1729 protein,			13.60
	127500	AW971353	Hs.162115	ESTs	11.20		
80	127524	AJ243596	Hs.94830	ESTs, Moderately similar to T03034 A-kin		7.80	
	127540	N45572	Hs.105362	Homo sapiens, clone MGC:18257, mRNA, com	3.53		
	127599	AA613204	Hs.150399	ESTs			13.80
	127609	X80031	Hs.530	collagen, type IV, alpha 3 (Goodpasture			28.00
	127662	W80755	Hs.8294	KIAA0196 gene product			19.80
85	127668	AJ343257	Hs.139993	ESTs			11.20

WO 02/086443

PCT/US02/12476

	127746	AI239495	Hs.120189	ESTs			14.18
	127812	AA741368	Hs.291434	ESTs	4.50		
	127817	AA835641	Hs.163085	ESTs			24.60
	127959	AJ302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L			9.20
5	127960	AI613226	Hs.41569	phosphatidic acid phosphatase type 2A			16.83
	127969	F06498	Hs.93748	Homo sapiens cDNA FLJ14676 fis, clone NT	13.60		
	128015	Z21169	Hs.334659	hypothetical protein MGC14139	7.00		
	128027	AI433721	Hs.164153	ESTs			37.40
	128077	AJ310330	Hs.128720	ESTs			9.60
10	128166	NM_008147	Hs.11801	interferon regulatory factor 6			9.24
	128226	AJ284940	Hs.289082	GM2 ganglioside activator protein	19.00		
	128305	AJ954968	Hs.279009	matrix Gla protein			10.40
	128341	AA191420	Hs.185030	ESTs	9.00		
	128527	AA504583	Hs.101047	transcription factor 3 (E2A immunoglobul			4.30
15	128539	R46163	Hs.258618	ESTs	12.60		
	128568	H12912	Hs.274691	adenylate kinase 3			4.56
	128572	AA933022	Hs.255583	interleukin enhancer binding factor 3, 9		16.80	10.00
	128777	AJ878918	Hs.10526	cysteine and glycine-rich protein 2			
	128781	N71826	Hs.105455	small nuclear ribonucleoprotein polypept			4.48
20	128796	AJ000152	Hs.105924	defensin, beta 2	8.12		
	128920	AA622037	Hs.166468	programmed cell death 5			4.62
	128924	BE279383	Hs.26557	plakophilin 3			4.04
	128971	H05132	Hs.107510	ESTs	12.60		
	129008	AL079648	Hs.301088	ESTs	8.80		
25	129041	BE382756	Hs.169902	solute carrier family 2 (facilitated glu			6.05
	129075	BE250162	Hs.63765	dihydrofolate reductase	2.59		
	129105	AJ769160	Hs.108681	Homo sapiens brain tumor associated prot		6.67	
	129189	AB023179	Hs.90959	KIAA0962 protein	8.00		
	129229	AF013758	Hs.109643	polyadenylate binding protein-interactin	4.00		
30	129241	AJ878857	Hs.109706	hematological and neurological expressed			4.06
	129300	W94197	Hs.110165	ribosomal protein L26 homolog	2.55		
	129404	AI267700	Hs.317584	ESTs	18.00		
	129457	X61959	Hs.207776	aspartylglucosaminidase	6.50		
	129466	L42583	Hs.334309	keratin 6A	12.94		
35	129494	AJ148976	Hs.112062	ESTs			11.00
	129505	AF061812	Hs.115947	keratin 16 (focal non-epidermolytic palm			4.46
	129541	AJ911527	Hs.11805	ESTs			12.00
	129565	AW163331	Hs.118778	KDEL (Lys-Asp-Glu-Leu) endoplasmic retic			4.70
	129703	BE388665	Hs.179999	Homo sapiens, clone IMAGE:3457003, mRNA			4.02
40	129720	AA156214	Hs.12152	APMCF1 protein			5.71
	129748	M16707	Hs.123053	H4 histone, family 2	3.50		
	129890	AI858872	Hs.282804	hypothetical protein FLJ22704			4.21
	129895	BE295568	Hs.13225	UDP-GalbetaGlcNAc beta 1,4- galactosylt	2.56		
	129945	BE514376	Hs.165998	PAI-1 mRNA-binding protein			4.03
45	130010	AA301116	Hs.142838	nucleolar phosphoprotein Nopp34		7.00	
	130026	T40480	Hs.332112	EST	6.40		
	130080	X14850	Hs.147097	H2A histone family, member X			4.65
	130149	AW067805	Hs.172665	methyltetrahydrofolate dehydrogenase	2.74		
50	130285	AA063546	Hs.75981	ubiquitin specific protease 14 (RNA-gua		7.40	
	130441	U63630	Hs.155637	protein kinase, DNA-activated, catalytic			3.91
	130482	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	4.87		
	130500	AB007913	Hs.158291	KIAA0444 protein			9.60
	130524	U89995	Hs.159234	forkhead box E1 (thyroid transcription f		13.40	
	130541	X05608	Hs.211584	neurofilament, light polypeptide (68kD)		8.20	
55	130553	AF062649	Hs.252587	pituitary tumor-transforming 1			6.06
	130567	AA383092	Hs.1608	replication protein A3 (14kD)		7.00	
	130577	M69241	Hs.162	insulin-like growth factor binding prote	3.04		
	130627	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	3.87		
60	130648	AI458165	Hs.17296	hypothetical protein MGC2376			16.20
	130697	L29472	Hs.1802	major histocompatibility complex, class			17.80
	130744	H59696	Hs.18747	POP7 (processing of precursor, S. cerevi			5.28
	130800	AJ187292	Hs.19574	hypothetical protein MGC5469			4.43
	130867	NM_001072	Hs.284239	UDP glycosyltransferase 1 family, polype	16.84		
65	130869	J03626	Hs.2057	uridine monophosphate synthetase (orotal			4.92
	130925	AF093419	Hs.169378	multiple PDZ domain protein			9.60
	130994	W17044	Hs.327337	ESTs		12.40	
	131028	AI879165	Hs.2227	CCAAT/enhancer binding protein (C/EBP),	10.21		
	131031	NM_001650	Hs.288650	aquaporin 4			9.80
	131041	T15767	Hs.22452	Homo sapiens mRNA for KIAA1737 protein,			9.60
70	131058	W28545	Hs.101514	hypothetical protein FLJ10342			17.00
	131090	AI143139	Hs.2268	visinin-like 1	2.74		
	131112	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f		8.80	
	131148	AW953575	Hs.303125	p53-induced protein PIGPC1	3.12		
	131185	BE280074	Hs.23950	cyclin B1	3.07		
75	131200	BE540516	Hs.293732	hypothetical protein MGC3195	3.07		
	131219	W25005	Hs.24395	small inducible cytokine subfamily B (Cy	2.87		
	131257	AW339037	Hs.24908	ESTs			14.67
	131375	AW293165	Hs.143134	ESTs		19.20	
80	131460	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	3.50		
	131476	AI521663	Hs.334644	hypothetical protein FLJ14668	15.00		
	131510	BE245374	Hs.27842	hypothetical protein FLJ11210		7.80	
	131646	BE302484	Hs.30057	MRS2 (S. cerevisiae)-like, magnesium hom		7.00	
	131786	BE000971	Hs.306083	Novel human gene mapping to chromosome 22	2.65		
	131839	AB014533	Hs.33010	KIAA0633 protein			35.20
85	131843	AA192315	Hs.184062	putative Rab5-interacting protein			4.11

WO 02/086443

PCT/US02/12476

	131877	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	19.00			
	131885	BE502341	Hs.3402	ESTs	6.48			
	131921	AA456093	Hs.34720	ESTs		8.40		
5	131945	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	56.00			
	131958	NM_014062	Hs.3556	ART-4 protein			3.82	
	131965	W79283	Hs.35962	ESTs	3.03			
	132000	AW247017	Hs.35978	melanoma antigen, family A, 3		9.60		
	132040	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	3.30			
10	132109	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	21.00			
	132114	NM_006152	Hs.40202	lymphoid-restricted membrane protein		8.40		
	132162	AA315805	Hs.94560	desmoglein 2			12.25	
	132164	AJ752235	Hs.41270	procollagen-lysine, 2-oxoglutarate 5-dio	2.70			
	132180	NM_004460	Hs.418	fibroblast activation protein, alpha	2.71			
15	132181	AW961231	Hs.16773	Homo sapiens clone TCCCA00427 mRNA sequ	3.83			
	132182	NM_014210	Hs.70499	ecotropic viral integration site 2A			13.20	
	132231	AA662910	Hs.42635	hypothetical protein DKFZp434K2435	9.50			
	132277	AK001745	Hs.184628	hypothetical protein FLJ10883	4.50			
	132328	NM_014787	Hs.44896	DnaJ (Hsp40) homolog, subfamily B, membe			9.20	
20	132394	AK001680	Hs.30488	DKFZP434F091 protein			19.80	
	132424	AA417878	Hs.48401	ESTs, Moderately similar to ALU8_HUMAN A		8.60		
	132528	T78736	Hs.50758	SMC4 (structural maintenance of chromoso		27.40		
	132543	BE568452	Hs.5101	protein regulator of cytokinesis 1	4.38			
	132544	L19778	Hs.51011	H2A histone family, member P		7.00		
25	132550	AW969253	Hs.170195	bone morphogenetic protein 7 (osteogenic	2.64			
	132552	BE621985	Hs.296922	thiopurine S-methyltransferase			15.83	
	132581	AK000631	Hs.52256	hypothetical protein FLJ20624		6.60		
	132617	AF037335	Hs.5338	carbonic anhydrase XII	4.95			
	132638	AJ796870	Hs.54277	DNA segment on chromosome X (unique) 992		8.20		
30	132653	Z15008	Hs.54451	laminin, gamma 2 (nicotin (100kD), kalini	4.38			
	132669	W38586	Hs.293981	guanine nucleotide binding protein (G pr			4.36	
	132710	W74001	Hs.55279	serine (or cysteine) proteinase inhibito	4.60			
	132771	Y10275	Hs.56407	phosphoserine phosphatase	3.71			
	132799	W73311	Hs.169407	SAC2 (suppressor of actin mutations 2,			9.48	
35	132833	U78525	Hs.57783	eukaryotic translation initiation factor			5.83	
	132892	AW834050	Hs.9973	tensin			12.00	
	132906	BE613337	Hs.234896	geminin	3.09			
	132959	AW014195	Hs.61472	ESTs, Weakly similar to YAE6_YEAST HYPOT			3.87	
	132962	AA576635	Hs.6153	CGI-48 protein	3.50			
40	132990	X77343	Hs.334334	transcription factor AP-2 alpha (activat	6.18			
	132994	AA112748	Hs.279905	clone HQ0310 PRO0310p1	3.19			
	133000	AL042444	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast	2.95			
	133050	X73424	Hs.63788	propionyl Coenzyme A carboxylase, beta p	2.55			
	133083	BE244588	Hs.6456	chaperonin containing TCP1, subunit 2 (b			4.00	
	133086	L17131	Hs.139800	high-mobility group (nonhistone chromoso			8.96	
45	133134	AF198620	Hs.65648	RNA binding motif protein 8A			4.28	
	133155	M58583	Hs.662	cerebellin 1 precursor			10.80	
	133181	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	3.00			
	133204	BE267695	Hs.254105	enolase 1, (alpha)			4.63	
50	133412	U41493	Hs.73112	guanine nucleotide binding protein (G pr		12.50		
	133421	AF134160	Hs.7327	claudin 1	2.85			
	133451	AW970026	Hs.73818	ubiquinol-cytochrome c reductase hinge p		6.80		
	133453	AJ659306	Hs.73826	protein tyrosine phosphatase, non-recept			4.66	
	133504	NM_004415	Hs.74316	desmoplakin (DPI, DPII)	6.14			
	133506	BE562958	Hs.74346	hypothetical protein MGC14353			4.55	
55	133615	M62843	Hs.75236	ELAV (embryonic lethal, abnormal vision,			17.80	
	133627	NM_002047	Hs.75280	glycyl-tRNA synthetase			4.85	
	133649	U25849	Hs.75393	acid phosphatase 1, soluble			6.34	
	133669	NM_006925	Hs.166975	splicing factor, arginine/serine-rich 5			14.00	
	133749	L20852	Hs.10018	solute carrier family 20 (phosphate tran		6.11		
60	133776	BE268649	Hs.177766	ADP-ribosyltransferase (NAD+; poly (ADP-			4.91	
	133865	AB011155	Hs.170290	discs, large (Drosophila) homolog 5	3.07			
	133946	AJ001258	Hs.173878	NIPSNAP, C. elegans, homolog 1			4.60	
	133973	N55540	Hs.78026	ESTs, Weakly similar to similar to ankyr			13.00	
	134047	BE262529	Hs.78771	phosphoglycerate kinase 1			3.85	
65	134098	BE513171	Hs.79086	mitochondrial ribosomal protein L3	2.56			
	134107	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte		8.20		
	134112	AW449809	Hs.79150	chaperonin containing TCP1, subunit 4 (d			4.08	
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	31.00			
70	134160	T98152	Hs.79432	fibrillin 2 (congenital contractual era		24.60		
	134168	AA398908	Hs.181634	Homo sapiens cDNA: FLJ23602 fis, clone L			6.71	
	134185	AA285136	Hs.301914	neuronal specific transcription factor D			14.74	
	134201	L35035	Hs.79886	ribose 5-phosphate isomerase A (ribose 5		8.40		
	134272	X76040	Hs.278614	protease, serine, 15	4.50			
75	134276	BE083936	Hs.80976	antigen identified by monoclonal antibod		9.00		
	134353	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m			16.40	
	134367	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	2.80			
	134380	AU077143	Hs.179555	minichromosome maintenance deficient (S.	4.68			
	134423	J53497	Hs.83006	CGI-139 protein			3.84	
	134469	AA279661	Hs.83753	small nuclear ribonucleoprotein polypept			5.81	
80	134470	X54942	Hs.83758	CDC28 protein kinase 2			4.21	
	134498	AW246273	Hs.84131	threonyl-tRNA synthetase			7.30	
	134502	BE148534	Hs.84168	UV-B repressed sequence, HUR 7		13.60		
	134510	NM_002757	Hs.250870	mitogen-activated protein kinase kinase			9.70	
	134548	N95406	Hs.333495	Deleted in split-hand/split-foot 1 regio			4.63	
85	134554	AK001741	Hs.8739	hypothetical protein FLJ10879	6.00			

WO 02/086443

PCT/US02/12476

5	134724	AF045239	Hs.321576	ring finger protein 22		12.00	
	134743	AA044163	Hs.89463	potassium large conductance calcium-acti	4.00		
	134781	AA374372	Hs.89626	parathyroid hormone-like hormone		25.20	
	134805	AD001528	Hs.89718	spermine synthase			4.58
	134853	BE268326	Hs.90280	5-aminimidazole-4-carboxamide ribonucle			4.79
	134859	D26488	Hs.90315	KIAA0007 protein		6.20	
	134891	RS1083	Hs.90787	ESTs		7.40	
	134960	BE246400	Hs.285176	acetyl-Coenzyme A transporter	4.00		
	134993	BE409809	Hs.301005	purine-rich element binding protein B			4.48
10	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	9.50		
	135080	AI761180	Hs.94211	rcd1 (required for cell differentiation,	5.00		
	135103	NM_003428	Hs.9450	zinc finger protein 84 (HPF2)		11.00	
	135145	AW014729	Hs.95262	nuclear factor related to kappa B bindin			4.01
	135184	U13222	Hs.96028	forkhead box D1		7.00	
15	135242	AI583187	Hs.9700	cyclin E1	13.50		
	135286	AW023482	Hs.97849	ESTs	6.46		
	135289	AW372589	Hs.9788	hypothetical protein MGC10924 similar to		8.80	
	135355	AK001652	Hs.99423	ATP-dependent RNA helicase	10.00		
	135371	NM_006025	Hs.997	protease, serine, 22	8.00		
20	135393	L11244	Hs.99886	complement component 4-binding protein,		14.60	

TABLE 5B shows the accession numbers for those primekeys lacking unigenelD's for Table 5A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
35	117079	1621717_1 H92325 T97125
	124305	242183_1 AW963221 AA344870 AA344871 H93331
	101502	18202_6 M26958
	109792	754958_1 R49625 F10674
	126034	1598157_1 H60340 N91637
40	102768	44641_1 U82321 H66077
	126345	1653833_1 N49713 N49819 W03810
	127066	1703458_1 R25066 R20144 R20145 Z43845
	127099	244301_1 AA347668 AW956810 Z44271 F07065 F07064 R13506
	119243	1774795_1 T12603 T12604
	125875	1566433_1 H14480 N98295
45	112054	1538292_1 R43590 F10439
	126979	171411_1 AA210954 AA211007
	126992	880655_1 AJ809521 H12174 Z42556
	122318	292419_1 AA429743 AA442754
	114699	135322_1 AA127386 R15644 AA127404
50	114793	150742_1 AA158245 AA158235
	108305	111550_1 AA071391 AA069892 AA069891
	108393	113411_1 AA075211 AA075245 AA075126 AA074946
	100867	tigr_HT4586 U14622
55	123731	genbank_AA609839 AA609839
	109700	genbank_F09609 F09609
	120715	genbank_AA292700 AA292700
	113702	genbank_T97307 T97307
	115113	genbank_AA256460 AA256460
60	101045	entrez_J05614 J05614
	108554	genbank_AA084948 AA084948
	108573	genbank_AA086005 AA086005
	119052	149538_1 R10889 R10888
	126522	416020_1 W31912 AI167491
65	126605	439280_1 AA676910 AA778853 AA778865 W68800
	103768	46922_1 W42667 AI580740 AI690440 AI561350 AW467906 AW151450 AI825927 AL041716 AI885600 AI742213 AW248624 AI955498 AA033947
		AA845593 AI623711 N68583 C00064 AA193567 AW083868 AW163216 AA191595 AA522778 AI628008 AI915518 AA843508 AI926195
		AA176265 AW167963 AA992115 W93647 AW103572 AI862994 AI342059 AA911719 AA176155 AA024712 AA069988 AA205591 AI591107
		AI199573 AI811766 AI275832 AI422233 AI191852 AI096682 AI580124 AI683612 AA582453 AA927559 AA486415 T32414 AI084978 H44849
		H44848 H20477 T91695 W47039 AA070055 AA024795 AA328855 AA379248 AA379330 AA385580 W25920 W03688 AA448359 AA093881
70		AW362477 AA089997 AI350255 W93479 N99688 AA932257 AW351459 H68590 AA663402 AA069771 AW087986 AI858420 AA600214
		AI970774 AI857712 AI683081 AI885584 AW131150 AI567981 AW002714 AW189973 AW075495 AW168303 AA953714 AW516881 AI357375
		AI566663 AW512676 AI570580 AI023690 AA448216 AI079853 AI422707 AA779516 AW026972 AW130082 AW162307 AW438646 AA709332
		AW192394 AI167350 AI217879 AI123152 AA719509 AI350480 AA663418 AI003634 AW118546 AA180261 AA442833 AI268625 AA888881
		AI038759 AA846723 AI248770 AA993694 AI280335 AI885107 AW518649 AA641563 AA95835 AA582521 AI276744 AA436478 AI017360
75		AI620763 AI859887 N73926 AI076327 AI741615 AI160617 AW172819 AI492005 AA677429 AA996334 AI693771 AI950039 AI456229 AI288515
		AI866186 T93293 AA173262 AA599779 AI680092 AW439316 AI084555 AI272672 AI583507 AW473219 AA738132 AW473283 AI367492
		AA995410 AI689624 AA206353 AI033095 AI040382 AA873630 AI221074 AI934840 AI416680 AA844306 R94503 AA773520 AA843169
		AA219425 AA629558 AI811719 AW411275 AI590981 W37907 AI591178 AI684051 AA983238 AA669347 AA976239 AA704570 AI628339
		AI884391 AI241580 AI003539 AW176687 AA009650 N34566 AI333493 AI186070 AA070827 AA411683 AI280884 AA872023 AA207255
80		AA021576 N71953 AI885888 AW076039 T15777 AI537673 AW248048 H09554 W93480 W47001 AW079114 AA063160 AA757453 R60788
		AI859431 H20478 AA218882 AA757465 AA100995 AI864135 AI934209 AA070503 H47008 AA219646 W61039 W93907 AW385050 W37967
		W78028 AA189007 AA479136 R93650 AA442312 T30287 AA847628 AA180262 AA009649 C03892 AW149464 AA310963 AA219693
		AA069747 R29207 AA094784 AA293615 AA447848 AI984167 N90393 C05097 N56499 AW292351 AW149681 AW473258 AA629322 AI004409
85		AW105577 AI954937 AI811070 AA902422 AW514437 AA535460 AA916877 AW517122 AA974657 AA975649 AW517130 AW517129 F31737
		W07688 AA193645 AA378994 AA489273 F32267 W39303 AA021181 N86810 AA406524 AA062553 AA436801 H08985 H15979 N40310

AA435789 AA232172 AW360778 W25862 R60282 AA436530 AA378894 AA187461 AI940535 AA604210 AA089514 AA360421 N88243 N84281
 AA209340 N56174 N88374 AA191088 AW247691 AA249013 AA093111 AA972536 AW298594 AA375893 T12139 W28186 AW243849
 AI288629 AA843996 W15260 AI188285 AW248079 R15836

5 119599 genbank_W45552 W45552
 112382 genbank_R59904 R59904
 105264 genbank_AA227934 AA227934
 100071 entrez_A28102 A28102
 123315 714071_1 AA496369 AA496646

10

Table 6A shows 99 genes up-regulated nonsmokers with lung cancer relative to smokers with lung cancer. These genes were selected from 59580 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

15

Pkey: Unique Eos probeset identifier number
 ExAccon: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigene number
 Unigene Title: Unigene gene title

20

R1: average of AI for samples from non-smokers with adenocarcinoma divided by the 90th percentile of AI for samples from smokers with adenocarcinoma
 R2: average of AI for samples from non-smokers with squamous cell carcinoma divided by the 90th percentile of AI for samples from smokers with squamous cell carcinoma

	Pkey	ExAccon	UnigenelD	Unigene Title	R1	R2
25	100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte		3.64
	101174	L17330	Hs.280	pre-TNK cell associated protein	15.00	
	101296	Y12490	Hs.85092	thyroid hormone receptor interactor 11		2.46
	101304	AA001021	Hs.6685	thyroid hormone receptor interactor 8		12.00
	101806	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias)		2.68
30	101972	S82472		gb:bata-poi=DNA polymerase beta (exon a		2.11
	102274	U30930	Hs.158540	UDP glycosyltransferase 8 (UDP-galactose	7.50	
	102394	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	7.50	
	102832	U92015		gb:Human clone 143789 defective mariner	13.50	
	103010	X52509	Hs.161640	tyrosine aminotransferase	9.50	
35	103439	X98266		gb:Hsapiens mRNA for ligase like protel		2.50
	103563	L02911	Hs.150402	activin A receptor, type I	9.00	
	103857	AI076795	Hs.45033	lacrimal proline rich protein		3.94
	104239	AB002367	Hs.21355	doublecortin and CaM kinase-like 1	13.50	
	104590	AW373062	Hs.83623	nuclear receptor subfamily 1, group I, m		12.66
40	104907	AA055829	Hs.196701	ESTs, Weakly similar to ALU1_HUMAN ALU	16.50	
	106131	BE514788	Hs.296244	SNARE protein		2.17
	106672	H47233	Hs.30643	ESTs	7.00	
	106872	T56887	Hs.18282	KIAA1134 protein	11.50	
	106960	AA156238	Hs.32501	ESTs		2.38
45	106971	Z43846	Hs.194478	Homo sapiens mRNA; cDNA DKFZp43401572 (f	9.50	
	107982	AA035375	Hs.57887	ESTs, Weakly similar to KIAA0758 protei		2.95
	108562	AA100798		gb:zmm26c06.s1 Stralagene pancreas (93720	16.50	
	108599	AB018549	Hs.69328	MD-2 protein	13.00	
	108663	BE219231	Hs.292653	ESTs, Weakly similar to T26845 hypotheti		2.40
50	109247	AA314907	Hs.85950	ESTs	7.00	
	109630	R44607	Hs.22672	ESTs		5.00
	110193	AI004874	Hs.310764	Homo sapiens mRNA; cDNA DKFZp434M082 (fr	12.50	
	110234	H24458	Hs.32085	EST	16.50	
	110644	R34207	Hs.268989	ESTs, Highly similar to type II CALM/AF1	8.00	
55	110886	AW274992	Hs.72249	three-PDZ containing protein similar to	17.00	
	111057	T79639	Hs.14629	ESTs	16.50	
	111950	AF071594	Hs.110457	Wolf-Hirschhorn syndrome candidate 1	11.00	
	112291	R53972	Hs.26026	ESTs		3.00
	112956	Z43784	Hs.75893	ankyrin 3, node of Ranvier (ankyrin G)		2.79
60	113009	T23699	Hs.7246	ESTs		4.50
	113060	BE564162	Hs.250820	hypothetical protein FLJ14827	9.79	
	113073	N39342	Hs.103042	microtubule-associated protein 1B	32.50	
	113074	AK001335	Hs.31137	protein tyrosine phosphatase, receptor t		3.82
	113121	T48011	Hs.8764	EST		2.21
65	113125	AA968672	Hs.8929	hypothetical protein FLJ11362	19.50	
	113757	AA703095	Hs.18631	ESTs		2.65
	113848	W52854	Hs.27099	hypothetical protein FLJ23293 similar to	6.00	
	113884	AI333076	Hs.28529	chromosome 12 open reading frame 2		6.00
	113936	W17056	Hs.83623	nuclear receptor subfamily 1, group I, m		4.63
70	114875	AA235609	Hs.236443	Homo sapiens mRNA; cDNA DKFZp564N1063 (7.00
	114987	AA251016	Hs.87808	EST		6.00
	115460	AW958439	Hs.38613	ESTs		2.27
	115722	W91892	Hs.59609	ESTs		9.00
	116261	AA481788	Hs.190150	ESTs	9.50	
75	116830	H61037	Hs.70404	ESTs, Weakly similar to ALU2_HUMAN ALU	8.50	
	116970	AB023179	Hs.9059	KIAA0962 protein	7.50	
	117178	H98675	Hs.269034	ESTs		2.68
	117757	AF088019	Hs.46732	EST	7.50	
	118283	AA287747	Hs.173012	ESTs, Weakly similar to A46010 X-linked	16.50	
80	118384	AF217525	Hs.49002	Down syndrome cell adhesion molecule		2.50
	118657	AI822106	Hs.49902	ESTs		2.39
	120328	AA923278	Hs.290905	ESTs, Weakly similar to protease [Hsapi		3.50
	120404	AB023230	Hs.96427	KIAA1013 protein	7.00	
	120524	AA261852	Hs.192905	ESTs	6.00	
85	120688	AW207555	Hs.97093	Homo sapiens cDNA: FLJ23004 fs, clone L	17.92	

	121558	AA412497		gb:z195g12s1 Soares_testis_NHT Homo sap		2.95
	121676	H55037	Hs.108145	ESTs	10.00	
	121836	AI024500	Hs.98612	ESTs	15.00	
	121938	AA426559	Hs.98610	ESTs	14.00	
5	122177	AA435789	Hs.98833	EST	8.93	
	123442	AA299652	Hs.111496	Homo sapiens cDNA FLJ11643 fis, clone HE	13.04	
	123551	AA608837		gb:af03h12s1 Soares_testis_NHT Homo sap	11.50	
	123756	AA609971	Hs.112795	EST	11.00	
	123861	AA620840		gb:af89g01.s1 Soares_testis_NHT Homo sap		2.50
10	124371	N24924	Hs.188601	ESTs	6.50	
	127477	BE328720	Hs.280651	ESTs		4.33
	127591	AI190540	Hs.131092	ESTs		3.02
	128252	AA455924	Hs.192228	ESTs	7.00	
	128426	AI265784	Hs.145197	ESTs		2.08
15	128925	R67419	Hs.21851	Homo sapiens cDNA FLJ12900 fis, clone NT		2.11
	128945	AI990506	Hs.8077	Homo sapiens mRNA; cDNA DKFZp547E184 (fr	10.00	
	129105	AI769160	Hs.108681	Homo sapiens brain tumor associated prot	15.50	
	129235	AW977238	Hs.126084	KIAA1055 protein		4.25
	129506	AB020684	Hs.11217	KIAA0877 protein	6.50	
20	129595	U09550	Hs.1154	oviductal glycoprotein 1, 120kD (mucin 9		10.00
	130160	AA305688	Hs.267695	UDP-Gal:betaGlcNAc beta 1,3-galactosyltr	20.00	
	130340	D82326	Hs.239106	solute carrier family 3 (cystine, dibasi	11.50	
	131220	AB023194	Hs.300855	KIAA0977 protein	17.50	
	131430	AI879148	Hs.26770	fatty acid binding protein 7, brain	6.10	
25	132114	NM_006152	Hs.40202	lymphoid-restricted membrane protein		6.15
	132458	AA935315	Hs.48965	Homo sapiens cDNA: FLJ21693 fis, clone C		5.58
	132647	NM_006927	Hs.54432	sialyltransferase 4B (beta-galactosidase	7.50	
	132655	D49372	Hs.54460	small inducible cytokine subfamily A (Cy		2.53
	132682	AI077500	Hs.54900	serologically defined colon cancer antig		2.50
30	132747	AA345241	Hs.55950	ESTs, Weakly similar to KIAA1330 protein		2.83
	132812	R50333	Hs.92186	Leman coiled-coil protein		3.82
	133337	AF085983	Hs.293676	ESTs		5.00
	133876	AL134906	Hs.771	phosphorylase, glycogen; liver (Hers dis		3.00
	134119	AW157837	Hs.79226	fasciculation and elongation protein zet		2.05
35	134464	AA302983	Hs.239720	CCR4-NOT transcription complex, subunit		2.27
	134542	M14156	Hs.85112	insulin-like growth factor 1 (somatomedi		11.50
	135002	AA448542	Hs.251677	G antigen 7B	87.00	
	135305	AA203555	Hs.98288	Homo sapiens cDNA FLJ14903 fis, clone PL		6.50

TABLE 6B show the accession numbers for those primekeys lacking unigenalIDs for Table 6A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

50

Pkey CAT number Accessions

55

108562	36375_1	AA100796 AF020589 AA074629 AA075946 AA100849 AA085347 AA126309 AA079311 AA079323 AA085274
103439	35330_1	X98266 N41124
123551	genbank_AA608837	AA608837
123861	genbank_AA620840	AA620840
102832	entrez_U92015	U92015
101972	entrez_S82472	S82472
121558	genbank_AA412497	AA412497

60

Table 7A shows 98 genes down-regulated in non-smokers with lung cancer relative to smokers with lung cancer. These genes were selected from 59680 probesets on the Eos/Altymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number				
	ExAccn:	Exemplar Accession number, Genbank accession number				
	UnigenelDt:	Unigene number				
	Unigena Title:	Unigene gene title				
10	R1:	90th percentile of AI for samples from smokers with adenocarcinoma divided by the average of AI for samples from non-smokers with adenocarcinoma.				
	R2:	90th percentile of AI for samples from smokers with squamous cell carcinoma divided by the average of AI for samples from non-smokers with squamous cell carcinoma.				
	Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2
15	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3		164.10
	100380	D82343	Hs.18551	neuroblastoma (nerve tissue) protein		77.40
	100576	X00356	Hs.37058	calclonin/calclonin-related polypeptid	102.40	
	100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte	463.80	
	101045	K01160		(NONE)	672.00	
20	101056	AW970254	Hs.889	Charot-Leyden crystal protein	66.00	
	101175	U82671	Hs.36980	melanoma antigen, family A, 2		77.20
	101497	W05150	Hs.37034	homeo box A5	62.80	
	101663	NM_003528	Hs.2178	H2B histone family, member Q	78.00	
	101677	NM_000715	Hs.1012	complement component 4-binding protein,	186.20	
25	101745	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	80.08	
	101941	S77583		gb:HERVK10/HUMMTV reverse transcriptase	99.20	
	102125	NM_006456	Hs.288215	sialyltransferase		103.10
	102242	U27185	Hs.82547	retinoic acid receptor responder (tazaro	67.00	
	102340	U37055	Hs.276657	macrophage stimulating 1 (hepatocyte gro	71.60	
30	102369	U39840	Hs.299867	hepatocyte nuclear factor 3, alpha		69.70
	102457	NM_001394	Hs.2359	dual specificity phosphatase 4	153.00	
	102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2		65.70
	102796	AL079646	Hs.107019	sympleskin; Huntingtin interacting protei		58.80
	102829	NM_006183	Hs.80962	neurotensin		268.80
35	103207	X72790		gb:Human endogenous retrovirus mRNA for	70.00	
	103242	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o		212.10
	103260	X78416	Hs.3155	casein, alpha		130.70
	103351	X89211		gb:H.sapiens DNA for endogenous retrovir	64.60	
40	104212	AB002238	Hs.173035	KIAA0300 protein	66.80	
	104252	AF002246	Hs.210863	cell adhesion molecule with homology to	63.80	
	104258	AF007216	Hs.5462	solute carrier family 4, sodium bicarbon	94.40	
	105024	AA126311	Hs.9879	ESTs	68.20	
	106260	AI097144	Hs.5250	ESTs, Weakly similar to ALU1_HUMAN ALU S		74.60
45	106440	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub		71.10
	106566	BE298210		gb:601118016F1 NIH_MGC_17 Homo sapiens c	73.20	
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr	83.80	
	106614	AA648459	Hs.335951	hypothetical protein AF301222		62.30
	106654	AW075485	Hs.286049	phosphoserine aminotransferase		202.40
	106999	H93281	Hs.10710	hypothetical protein FLJ20417		89.60
50	108700	AA121518	Hs.193540	ESTs, Moderately similar to 2109260A B c		66.40
	108810	AW295647	Hs.71331	hypothetical protein MGC5350		95.50
	108857	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act		63.40
	109697	AA989362	Hs.293780	ESTs	85.00	
55	109691	T65568	Hs.12860	ESTs		58.70
	109704	AI743880	Hs.12876	ESTs		60.60
	110942	R63503	Hs.28419	ESTs	76.40	
	111722	R23924	Hs.23596	EST	74.60	
	112891	T03927	Hs.293147	ESTs, Moderately similar to A46010 X-li	64.80	
60	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f		76.70
	113073	N39342	Hs.103042	microtubule-associated protein 1B		120.20
	114251	H15261	Hs.21948	ESTs	127.20	
	115230	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	174.00	
	115291	BE545072	Hs.122579	hypothetical protein FLJ10461		91.00
65	115815	AW905328	Hs.180842	ribosomal protein L13	66.40	
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH		226.60
	115965	AA001732	Hs.173233	hypothetical protein FLJ10970	82.80	
	116107	AL133916	Hs.172572	hypothetical protein FLJ20093		361.60
	116552	D20508	Hs.164649	hypothetical protein DKFZp434H247	69.00	
70	116571	D45652		gb:HUMGS02848 Human adult lung 3' direct	64.20	
	118465	N66741		gb:yz33g08.s1 Morton Fetal Cochlea Homo		63.50
	120484	AA253170	Hs.96473	EST	81.60	
	120983	AA398209	Hs.97587	EST		81.10
	121034	AL389951	Hs.271623	nucleoporin 50kD		66.20
75	121423	AW973352	Hs.290585	ESTs	64.40	
	122553	AA451884	Hs.190121	ESTs		60.40
	122946	AI718702	Hs.308026	major histocompatibility complex, class	188.60	
	123130	AA487200		gb:ab19f02.s1 Stratagene lung (937210) H		80.20
	124472	N52517	Hs.102670	EST	71.00	
80	124526	N52096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci		104.90
	125489	H49193	Hs.124984	ESTs, Moderately similar to ALLU7_HUMAN A		72.00
	125731	R61771	Hs.26912	ESTs		69.90
	125747	NM_002884	Hs.865	RAP1A, member of RAS oncogene family	69.00	
	126020	H79863	Hs.114243	ESTs		62.40
	126547	U47732	Hs.84072	transmembrane 4 superfamily member 3		62.80
85	126966	R38438	Hs.182575	solute carrier family 15 (H+/peptide tra		60.10

	127472	AA761378	Hs.192013	ESTs	70.20	
	127610	AA950867	Hs.150271	ESTs, Highly similar to unnamed protein	64.00	
	127742	AW293496	Hs.180138	ESTs	85.20	
	127987	AJ022103	Hs.124511	ESTs	95.60	
5	128233	AW889132	Hs.11916	ribokinase		78.90
	128420	AA650274	Hs.41296	fibronectin leucine rich transmembrane p		106.90
	128766	AW160432	Hs.296460	craniofacial development protein 1	66.80	
	129014	AW935187	Hs.170162	KIAA1357 protein		58.53
	129215	AB040930	Hs.126085	KIAA1497 protein	64.20	
10	130090	H97878	Hs.132390	zinc finger protein 36 (KOX 16)	63.80	
	130385	AW067800	Hs.155223	stannocalcin 2		139.60
	130732	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)		64.60
	131025	AB040900	Hs.6189	KIAA1467 protein	64.40	
	131241	BE501914	Hs.24654	Homo sapiens cDNA FLJ11640 fis, clone HE	76.20	
15	131775	AB014548	Hs.31921	KIAA0648 protein	97.80	
	132240	AB018324	Hs.42676	KIAA0781 protein		71.00
	132656	NM_001448	Hs.58367	glypican 4		88.40
	132977	AA093322	Hs.301404	RNA binding motif protein 3	133.20	
	133749	L20852	Hs.10018	solute carrier family 20 (phosphate tran		59.30
20	133818	AJ110684	Hs.7645	fibrinogen, B beta polypeptide	341.00	
	134264	AF149297	Hs.8087	NAG-5 protein		64.30
	134265	M83772	Hs.80876	flavin containing monooxygenase 3		232.53
	134346	X84002	Hs.82037	TATA box binding protein (TBP)-associate	66.00	
	134395	AA456539	Hs.8262	lysosomal-associated membrane protein 2		75.80
25	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su		108.30
	135056	N75765	Hs.93765	lipoma HM/GIC fusion partner	71.40	
	135309	A1564123	Hs.42500	ADP-ribosylation factor-like 5	70.40	

30 TABLE 7B shows the accession numbers for those primekeys lacking unigenelD's for Table 7A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

35 Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
103207	30635_4	X72790
106566	120358_1	BE298210 AJ672315 AW086489 BE298417 AA455921 AA902537 BE327124 R14963 AA085210 AW274273 AI333584 AI369742 AI039658 AI885095 AI476470 AJ287650 AI885299 AI985381 AW592624 AW340136 AJ268556 AA456390 AI310815 AA484951
116571	genbank_D45652	D45652
118466	genbank_N66741	N66741
101046	entrez_K01160 K01160	
101941	entrez_S77583 S77583	
103351	entrez_X89211 X89211	
123130	genbank_AA487200	AA487200

Table 8A shows 1720 genes either up or down-regulated in lung tumors or chronically diseased lung relative to a broad collection of over 40 distinct normal body tissues. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 39494 probesets on the Eos/Affymetrix Hu02 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5

Pkey: Unique Eos probeset identifier number
ExAccn: Exemplar Accession number, Genbank accession number
UnigeneID: Unigene number
Unigene Title: Unigene gene title

10

R1: 70th percentile of AI for lung tumors divided by 90th percentile of AI for normal lung
R2: 70th percentile of AI for chronically diseased lung divided by 90th percentile of AI for normal lung

	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
15	300097	AJ916973	Hs.213603	ESTs	5.46	4.69
	300117	AW189787	Hs.147474	ESTs	0.58	0.56
	300197	AJ686661	Hs.218286	ESTs	4.26	5.44
	300201	AJ308300		gb:ta90c06.x1 NCL_CGAP_Bm20 Homo sapien	0.62	0.83
20	300225	AJ989963	Hs.197505	ESTs	1.68	1.75
	300247	AW274682	Hs.161394	ESTs	1.08	2.28
	300256	AJ469095	Hs.298241	Transmembrane protease, serine 3	0.86	1.00
	300337	AJ707881	Hs.202090	ESTs	5.80	9.09
	300362	Z42308		gb:HSC0FB121 normalized infant brain cDN	4.18	12.78
	300374	AJ859947	Hs.314158	ESTs	2.99	4.38
25	300387	AW270150	Hs.254516	ESTs	1.50	2.53
	300440	AJ421541	Hs.146164	ESTs	3.98	5.25
	300441	R10367	Hs.307921	EST, Weakly similar to Z232_HUMAN ZINC F	3.18	6.80
	300449	AJ362967	Hs.132221	hypothetical protein FLJ12401	0.43	0.62
	300469	AW135830	Hs.233955	hypothetical protein FLJ20401	0.16	0.83
30	300552	X85711	Hs.21838	hypothetical protein FLJ11191	4.10	9.75
	300627	W27363		gb:ab37d01.r1 Stratagene HeLa cell s3 93	4.60	12.60
	300630	AW118822	Hs.128757	ESTs	2.91	5.86
	300716	AJ216113	Hs.126280	hypothetical protein FLJ23393	1.00	0.92
	300738	AJ623332	Hs.130541	KIAA1542 protein	1.82	1.71
35	300777	AA235361	Hs.96840	KIAA1527 protein	4.48	8.22
	300790	AJ492471	Hs.188270	ESTs	1.29	1.18
	300832	AJ688147	Hs.220615	ESTs, Weakly similar to T03829 transcript	5.51	8.56
	300836	Z44942	Hs.22958	calcium channel alpha2-delta3 subunit	4.90	6.34
	300838	AJ582897	Hs.192570	hypothetical protein FLJ22028	1.70	2.81
40	300878	AW449802	Hs.285901	Homo sapiens cDNA FLJ20428 fis, clone KA	4.56	7.91
	300897	AJ890356	Hs.127804	ESTs, Weakly similar to T17233 hypotheti	2.23	1.58
	300926	AA504860		gb:ab03a10.s1 Stratagene fetal retina 93	2.13	3.50
	300960	AJ041019	Hs.152454	ESTs	2.74	4.46
	300961	AW204069	Hs.312716	ESTs, Weakly similar to unnamed protein	1.00	1.00
45	300962	AA593373	Hs.293744	ESTs	1.46	1.51
	300967	AA565209	Hs.269439	ESTs	0.39	1.30
	300987	AW450840	Hs.148590	ESTs, Weakly similar to AF208846 1 BM-00	1.49	1.08
	300988	AJ927208	Hs.208952	ESTs	0.16	0.37
	301050	AW136973	Hs.288516	ESTs, Weakly similar to S69890 mitogen I	3.23	1.94
50	301098	AA677570	Hs.185918	ESTs	6.76	14.28
	301157	AA729905	Hs.231916	ESTs	3.16	8.85
	301162	AJ142118	Hs.129004	ESTs	1.68	7.18
	301170	AA737594	Hs.247606	ESTs	4.40	6.42
	301192	AJ808751	Hs.121188	ESTs	6.38	11.59
55	301193	AA758115	Hs.128350	ESTs, Weakly similar to JC5423 2-hydroxy	4.35	7.78
	301257	AW297762	Hs.255690	ESTs	1.56	1.61
	301281	AA843986	Hs.190586	ESTs	2.19	1.78
	301341	AJ819198	Hs.208229	ESTs	0.76	0.76
	301382	AA912839	Hs.163369	ESTs	1.00	1.81
60	301407	AW450466	Hs.126830	ESTs	1.48	1.51
	301452	AA975688	Hs.159955	ESTs	0.51	1.46
	301483	AW272467	Hs.254655	Unlabeled	2.40	5.02
	301494	AJ678034	Hs.131099	ESTs	2.79	3.41
	301521	AJ733621	Hs.133011	zinc finger protein 117 (HPF9)	0.67	0.67
65	301531	AJ077462	Hs.134084	ESTs	2.52	3.76
	301580	AJ878959	Hs.73737	splicing factor, arginine/serine-rich 1	7.41	11.92
	301676	Z43570	Hs.27453	ESTs, Moderately similar to G01251 Rar p	8.31	10.70
	301690	F05865	Hs.108323	ubiquitin-conjugating enzyme E2E 2 (homo	2.70	4.22
	301718	F07744	Hs.7987	DKFZP434F162 protein	4.20	8.78
70	301799	AA384252	Hs.286132	D15F37 (pseudogene)	5.93	7.04
	301804	AA581004	Hs.62180	anillin (Drosophila Scraps homolog), act	1.70	0.76
	301822	X17033	Hs.271986	Integrin, alpha 2 (CD49B, alpha 2 subuni	1.58	1.36
	301846	R20002	Hs.6823	hypothetical protein FLJ10430	1.00	1.00
	301868	T71508	Hs.13861	ESTs, Weakly similar to pH sensitive max	2.88	5.49
75	301882	T78054		gb:yc97g09.r1 Soares infant brain 1N1B H	2.28	3.80
	301905	AJ991127	Hs.117202	ESTs	1.00	1.00
	301948	AA344647	Hs.116724	aldo-keto reductase family 1, member B11	5.28	2.28
	301960	AW070252	Hs.27973	KIAA0874 protein	5.38	6.48
	302011	T91418	Hs.125156	transcriptional adaptor 2 (ADA2, yeast,	3.03	3.42
80	302016	N40834	Hs.23495	hypothetical protein FLJ11252	1.00	1.25
	302041	NM_001501	Hs.129715	gonadotropin-releasing hormone 2	0.71	0.99
	302072	AJ238381	Hs.132576	paired box gene 9	1.60	1.71
	302094	AJ286176	Hs.6786	ESTs	0.52	1.20
	302095	AW044300	Hs.137506	Homo sapiens BAC clone RP11-120J2 from 7	2.75	4.93
85	302148	AW269518	Hs.23244	ESTs	3.04	3.87

	302155	AJ088485	Hs.144759	ESTs	0.45	1.15
	302201	AJ006276	Hs.159003	transient receptor potential channel 6	0.33	0.84
	302202	AF097159	Hs.159140	UDP-Gal4betaGlcNAc 4- galactosylt	0.52	0.94
5	302206	AJ937193	Hs.41143	phosphoinositide-specific phospholipase	2.76	3.65
	302209	AF047445	Hs.159297	killer cell lectin-like receptor subfam	1.00	1.00
	302235	AL049987	Hs.166361	Homo sapiens mRNA; cDNA DKFZp554F112 (tr	1.68	1.50
	302290	AL117607	Hs.175563	Homo sapiens mRNA; cDNA DKFZp554N0763 (f	1.00	2.11
	302328	AA354849	Hs.23240	Homo sapiens cDNA FLJ13496 fis, clone PL	9.38	13.08
10	302346	AL039101	Hs.194625	dynein, cytoplasmic, light intermediate	3.27	7.24
	302360	AJ010901	Hs.198267	mucin 4, tracheobronchial	2.54	1.83
	302384	Y08982	Hs.202676	synaptonemal complex protein 2	1.00	0.91
	302406	U86751	Hs.211956	CD3-epsilon-associated protein; antisens	2.63	2.67
	302409	AF155156	Hs.218028	adaptor-related protein complex 4, epsil	5.82	9.34
	302423	AB028977	Hs.225974	KIAA1054 protein	3.66	3.18
15	302432	AL080068	Hs.272534	Homo sapiens mRNA; cDNA DKFZp554J062 (fr	2.44	6.77
	302435	AF092047	Hs.227277	sine oculis homeobox (Drosophila) homolo	0.44	0.84
	302437	AB024730	Hs.227473	UDP-N-acetylglucosaminase-1,3-D-mannosid	4.18	5.64
	302455	AA356923	Hs.240770	nuclear cap binding protein subunit 2, 2	1.85	0.92
20	302472	AA317451	Hs.6335	SW/SNF related, matrix associated, acti	2.04	2.13
	302476	AF182294	Hs.241578	U6 snRNA-associated Sm-like protein LSm8	1.44	1.89
	302489	T80660	Hs.230424	Homo sapiens cDNA FLJ13540 fis, clone PL	0.51	1.10
	302490	AA885502	Hs.187032	ESTs	2.64	4.87
	302562	AJ005585	Hs.48956	gap junction protein, beta 6 (connexin 3	5.34	2.68
25	302566	AA085996	Hs.248572	hypothetical protein FLJ22965	1.00	1.21
	302630	AB029488	Hs.272100	SMS3 protein	0.52	1.24
	302634	AB032953	Hs.173560	odd Ozifen-m homolog 2 (Drosophila, mous	1.00	1.00
	302638	AA463798	Hs.102696	MCT-1 protein	1.58	1.02
	302647	X57723	Hs.198273	NADH dehydrogenase (ubiquinone) 1 beta s	2.72	6.85
30	302655	AJ227892	Hs.146274	ESTs	1.00	4.32
	302656	AW293005	Hs.70704	Homo sapiens, clone IMAGE:2823731, mRNA,	2.97	0.93
	302668	AA580691	Hs.180789	S164 protein	0.80	0.95
	302679	H65022		gb:yu66g11.1 Weizmann Olfactory Epithel	1.68	5.04
	302680	AW192334	Hs.38218	ESTs	2.70	7.98
35	302697	AJ001408		gb:Homo sapiens mRNA for immunoglobulin	4.25	8.13
	302705	U09060		gb:Human immunoglobulin heavy chain, V-r	3.91	8.68
	302711	L08442		gb:Human autonomously replicating sequen	2.20	2.73
	302719	W69724	Hs.288959	hypothetical protein FLJ20920	0.54	1.02
	302742	L12069		gb:Homo sapiens (clone WR4.10VH) anti-th	7.28	11.57
40	302755	AW384815	Hs.149208	KIAA1555 protein	1.57	2.38
	302771	H98476	Hs.42522	ESTs	2.94	4.68
	302789	AJ245067		gb:Homo sapiens mRNA for immunoglobulin	3.49	6.31
	302795	AJ245313	Hs.272638	hypothetical protein FLJ10494	0.80	2.74
	302802	Y08250		gb:H.sapiens mRNA for variable region of	1.13	0.77
45	302803	AA442824	Hs.293951	ESTs, Moderately similar to putative DNA	3.14	10.68
	302812	N31301	Hs.152664	hypothetical protein FLJ20051	3.04	8.24
	302847	X98940		gb:H.sapiens rearranged Ig heavy chain (1.80	1.92
	302885	AL137763	Hs.132127	hypothetical protein LOC57822	1.00	1.00
	302943	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti	0.53	0.67
	302977	AW263124	Hs.315111	hypothetical protein FLJ12894	2.45	2.62
50	303006	AF078950	Hs.24139	Homo sapiens cDNA: FLJ23137 fis, clone L	4.88	8.61
	303011	AF090405		gb:Homo sapiens clone 2A1 scFV antibody	1.41	1.86
	303013	F07898	Hs.288958	RAB22A, member RAS oncogene family	1.51	1.19
	303061	AF151882	Hs.27693	peptidylprolyl isomerase (cyclophilin)-I	0.72	0.76
55	303077	AF163305		gb:H.sapiens T-cell receptor mRNA	1.17	3.90
	303090	AA443259	Hs.146286	kinesin family member 13A	4.08	6.46
	303091	AF192913	Hs.130683	zinc finger protein 180 (HHZ168)	2.50	4.37
	303094	AF195513	Hs.278953	Pur-gamma	5.38	8.38
	303095	AF202051	Hs.134079	NM23-H8	3.26	4.08
60	303131	AW081061	Hs.103180	DC2 protein	2.02	1.83
	303195	AA082211	Hs.233936	myosin, light polypeptide, regulatory, n	1.32	3.95
	303196	AA082298	Hs.59710	ESTs	0.77	0.53
	303216	AA581439	Hs.152328	ESTs	0.24	0.63
	303222	AA333538	Hs.204501	hypothetical protein FLJ10534	3.56	6.22
65	303234	AA132255	Hs.143951	ESTs	2.28	3.17
	303251	AW340037	Hs.115897	protocadherin 12	0.38	1.02
	303295	AA205625	Hs.208067	ESTs	2.30	1.00
	303297	T80072	Hs.13423	Homo sapiens clone 24468 mRNA sequence	1.86	4.48
	303316	AF033122	Hs.14125	p53 regulated PA26 nuclear protein	0.10	0.80
70	303467	AA398801	Hs.323397	ESTs	4.54	9.65
	303506	AA340605	Hs.105887	ESTs, Weakly similar to Homolog of rat Z	0.09	0.04
	303552	AA359799	Hs.224662	ESTs, Weakly similar to unnamed protein	1.00	1.72
	303598	AA382814		gb:EST96097 Testis I Homo sapiens cDNA 5	4.96	9.14
	303637	AF056083	Hs.24879	phosphatidic acid phosphatase type 2C	2.06	2.02
75	303655	AA504702	Hs.258802	ATPase, (Na+)/K+ transporting, beta 4 po	1.00	1.24
	303756	AJ738488	Hs.115838	ESTs	1.08	1.43
	303856	AA968589	Hs.180532	glucose phosphate isomerase	1.76	1.31
	303893	N88597	Hs.113503	karyopherin (importin) beta 3	2.30	2.57
	303907	AW467774	Hs.171880	polymerase (RNA) II (DNA directed) polyp	3.10	5.79
	303946	AW474196	Hs.306637	Homo sapiens cDNA FLJ12363 fis, clone MA	5.06	11.86
80	303978	AW513315		gb:cx43c12.x1 NCL_CGAP_U11 Homo sapiens	5.14	7.31
	303981	AW513804	Hs.278834	ESTs, Weakly similar to ALU1_HUMAN ALU S	2.83	4.06
	303990	AW515465		gb:cx71a11.x1 NCL_CGAP_Kd8 Homo sapiens	1.15	2.35
	303998	AW516449		gb:cx58f05.x1 NCL_CGAP_U12 Homo sapiens	2.20	9.35
	303999	AW516611		gb:cx70b11.x1 NCL_CGAP_Ov39 Homo sapiens	4.85	6.28
85	304006	AW517947		gb:cx55h02.x1 NCL_CGAP_U12 Homo sapiens	3.21	4.07

	304008	AW518198	Hs.3297	ribosomal protein S27a	6.50	11.08
	304009	AW518206	Hs.181165	eukaryotic translation elongation factor	1.88	3.27
	304024	T03036		gb:FB2187 Fetal brain, Stratagene Homo s	2.15	3.55
5	304026	T03160		gb:FB26F2 Fetal brain, Stratagene Homo s	5.88	11.80
	304028	T03266		gb:FB7C1 Fetal brain, Stratagene Homo sa	5.59	13.46
	304036	T16855	Hs.244621	ribosomal protein S14	6.55	14.43
	304046	T54803		gb:yb42d06.s1 Stratagene fetal spleen (9	6.18	12.19
	304061	T61521		gb:yb73g01.s1 Stratagene ovary (937217)	2.64	8.23
10	304063	T62536		gb:yc04c12.s1 Stratagene lung (937210) H	0.53	1.61
	304097	R25376	Hs.177592	ribosomal protein, large, P1	6.49	11.67
	304114	R78946		gb:yi87g02.s1 Soares placenta Nb2HP Homo	2.90	4.18
	304122	H28966		gb:ym31a06.s1 Soares infant brain 1N1B H	1.00	2.76
	304155	H68695		gb:yr78b05.s1 Soares fetal liver spleen	0.79	1.18
15	304203	N56929		gb:yy82d08.s1 Soares_multiple_sclerosis_	4.28	11.34
	304234	W81608		gb:zd68h06.s1 Soares_fetal_heart_NbHH19W	6.47	11.03
	304267	AA064862	Hs.73742	ribosomal protein, large, P0	1.34	1.16
	304270	AA069711	Hs.297753	vimentin	3.40	5.40
	304287	AA079286	Hs.78466	proteasome (prosome, macropain) 26S sub	2.93	4.42
20	304348	AA179868		gb:zp38g12.s1 Stratagene muscle S37209 H	3.98	10.96
	304415	AA290747	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	3.32	5.99
	304430	AA347682		gb:EST54044 Fetal heart II Homo sapiens	1.00	1.00
	304456	AA411240		gb:zv26g05.s1 Soares_NhHMPu_S1 Homo sapi	1.42	3.33
	304521	AA464716		gb:zx62c11.s1 Soares ovary tumor NbHOT H	2.18	1.15
25	304526	AA476427		gb:zz02c05.s1 Soares_total_fetus_Nb2HF8_	5.38	14.11
	304542	AA482602	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.16	8.23
	304546	AA486074	Hs.297681	serine (or cysteine) proteinase inhibitor	0.55	1.20
	304607	AA513322		gb:nh85e08.s1 NCI_CGAP_Br1.1 Homo sapien	1.95	2.10
	304640	AA524440	Hs.111334	ferritin, light polypeptide	2.10	2.83
30	304650	AA527489	Hs.3463	ribosomal protein S23	3.33	12.62
	304735	AA576453		gb:nm75h11.s1 NCI_CGAP_Co9 Homo sapiens	1.33	0.88
	304760	AA590401		gb:nm13g09.s1 NCI_CGAP_Co12 Homo sapiens	3.68	8.14
	304849	AA588157	Hs.13801	KIAA1685 protein	2.77	3.70
	304917	AA602685	Hs.284136	PRO2047 protein	7.16	11.01
35	304921	AA603092	Hs.297753	vimentin	2.47	4.24
	304966	AA613893	Hs.282435	ESTs	6.78	11.66
	304987	AA618044	Hs.300697	immunoglobulin heavy constant gamma 3 (G	0.90	1.23
	305016	AA626876		gb:zu89h06.s1 Soares_testis_NHT Homo sap	6.46	10.17
	305034	AA630128		gb:ab99c04.s1 Stratagene lung (937210) H	1.00	1.00
40	305072	AA641012		gb:nr72a12.s1 NCI_CGAP_Pr24 Homo sapiens	5.68	11.59
	305111	AA644187	Hs.303405	ESTs	1.48	1.37
	305148	AA654070		gb:nf01g08.s1 NCI_CGAP_Lym3 Homo sapiens	1.76	4.61
	305159	AA659166	Hs.275668	EST, Weakly similar to EF1D_HUMAN ELONG	1.00	2.15
	305190	AA665955		gb:ag57d12.s1 Gessler Wilms tumor Homo s	5.31	8.14
45	305232	AA670052	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	0.78	1.18
	305235	AA670480		gb:ag37e01.s1 Jia bone marrow stroma Hom	3.11	8.66
	305245	AA676695	Hs.81328	nuclear factor of kappa light polypeptid	4.38	7.53
	305312	AA700201		gb:zj44f07.s1 Soares_fetal_liver_spleen_	2.13	2.66
	305322	AA701597	Hs.163019	EST	1.20	1.40
50	305394	AA720942	Hs.300697	immunoglobulin heavy constant gamma 3 (G	1.16	0.68
	305413	AA724659		gb:al10f08.s1 Soares_parathyroid_tumor_N	5.86	9.87
	305447	AA737856		gb:nx10c08.s1 NCI_CGAP_GC3 Homo sapiens	2.21	2.86
	305476	AA745664	Hs.287445	hypothetical protein FLJ11726	3.36	6.54
	305483	AA748030	Hs.303512	EST	1.00	2.02
55	305528	AA769156		gb:nz12e05.s1 NCI_CGAP_GCB1 Homo sapiens	6.44	9.10
	305612	AA782347	Hs.272572	hemoglobin, alpha 2	0.19	0.79
	305614	AA782866		gb:aj09h02.s1 Soares_parathyroid_tumor_N	1.00	1.00
	305616	AA782884	Hs.275865	ribosomal protein S18	7.57	10.20
	305637	AA806124		gb:oe29a12.s1 NCI_CGAP_Pr25 Homo sapiens	4.78	12.42
60	305639	AA806138		gb:oe29c12.s1 NCI_CGAP_Pr25 Homo sapiens	0.89	0.70
	305650	AA807709		gb:nw31e04.s1 NCI_CGAP_GC80 Homo sapiens4.49		8.71
	305690	AA813477		gb:ai67a05.s1 Soares_testis_NHT Homo sap	4.91	9.40
	305726	AA828156	Hs.73742	ribosomal protein, large, P0	0.19	0.81
	305728	AA828209		gb:of34a02.s1 NCI_CGAP_Kd6 Homo sapiens	5.12	9.29
65	305759	AA835353		gb:ak72b06.s1 Barstead spleen HPLRB2 Hom	1.66	4.11
	305792	AA845256		gb:ak84a08.s1 Barstead spleen HPLRB2 Hom	2.34	4.25
	305864	AA854374	Hs.73742	ribosomal protein, large, P0	0.30	1.40
	305901	AA872958		gb:oh63h08.s1 NCI_CGAP_Kd5 Homo sapiens	2.10	5.21
	305910	AA875981		gb:nx21h02.s1 NCI_CGAP_GC3 Homo sapiens	0.32	1.01
70	306015	AA897116		gb:am08b07.s1 Soares_NFL_T_GBC_S1 Homo s1.56		1.12
	306017	AA897221	Hs.109058	ribosomal protein S6 kinase, 90kD, polyp	5.21	7.90
	306020	AA897630	Hs.130027	EST	1.96	6.59
	306063	AA906316		gb:ok03g03.s1 Soares_NFL_T_GBC_S1 Homo s	7.38	20.69
	306065	AA906725		gb:ok78g02.s1 NCI_CGAP_GC4 Homo sapiens	7.19	13.48
75	306104	AA910956		gb:ok85h11.s1 NCI_CGAP_Kd3 Homo sapiens	6.50	9.13
	306109	AA911861		gb:ox21a07.s1 NCI_CGAP_PNS1 Homo sapiens	4.21	5.25
	306148	AA917409	Hs.288036	IRNA isopentenylpyrophosphate transferas	2.20	2.70
	306242	AA932805		gb:oo60g04.s1 NCI_CGAP_Lu5 Homo sapiens	2.84	5.35
	306288	AA936900		gb:oi53h05.s1 NCI_CGAP_HN3 Homo sapiens	1.60	1.12
80	306325	AA953072	Hs.210546	interleukin 21 receptor	1.65	2.26
	306353	AA961382	Hs.275865	ribosomal protein S18	3.78	6.32
	306375	AA968650	Hs.276018	EST, Moderately similar to JC4662 ribos	4.30	5.74
	306396	AA970223		gb:op09d05.s1 NCI_CGAP_Kd6 Homo sapiens	0.95	2.45
	306428	AA975110	Hs.191228	hypothetical protein FLJ20284	3.19	4.10
85	306442	AA976899		gb:ox35e09.s1 NCI_CGAP_GC4 Homo sapiens	4.67	7.44
	306446	AA977348		gb:ox72e12.s1 NCI_CGAP_Kd6 Homo sapiens	3.92	6.27

	306458	AA978186		gb:op33c06.s1 Soares_NFL_T_GBC_S1 Homo s.	3.35	5.77
	306467	AA983508	Hs.163593	ribosomal protein L18a	3.72	5.37
	306510	AA988546		gb:or84d07.s1 NCL_CGAP_Lu5 Homo sapiens	1.00	1.00
5	306555	AA994304	Hs.276083	EST, Weakly similar to RL23_HUMAN 60S R	6.61	10.91
	306557	AA994530		gb:ou57e08.s1 NCL_CGAP_Br2 Homo sapiens	16.20	31.83
	306572	AA995686		gb:os25c12.s1 NCL_CGAP_Kid5 Homo sapiens	2.51	6.52
	306582	AA996248		gb:os18c10.s1 NCL_CGAP_Kid5 Homo sapiens	1.42	3.13
	306598	AI000320	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.91	8.68
10	306605	AI000497	Hs.119500	ribosomal protein, large P2	1.96	8.60
	306656	AI004024		gb:ou11b07.x1 Soares_NFL_T_GBC_S1 Homo s	0.11	0.45
	306676	AI005603	Hs.284136	PRO2047 protein	9.56	17.28
	306688	AI015615		gb:ov29f10.x1 Soares_testis_NHT Homo sap	1.86	3.60
	306702	AI022565	Hs.307670	EST	1.47	1.19
	306728	AI027359	Hs.272572	hemoglobin, alpha 2	1.28	2.83
15	306751	AI032589		gb:ow70h12.s1 Soares_fetal_liver_spleen_	3.91	5.21
	306767	AI038963	Hs.249118	ESTs	3.33	6.06
	306892	AI092465		gb:xq75h12.x1 Soares_fetal_heart_NbHH19W	3.77	7.46
	306897	AI093967		gb:xq33c06.s1 Soares_NhHMPu_S1 Homo sapi	2.12	2.85
20	306956	AI125111		gb:am66f03.s1 Barstead spleen HPLRB2 Hom	6.10	10.52
	306958	AI125152		gb:am55e09.x1 Johnston frontal cortex Ho	1.72	1.56
	307035	AI142774	Hs.119122	ribosomal protein L13a	2.00	4.70
	307041	AI144243		gb:xq85b12.x1 Soares_fetal_heart_NbHH19W	9.12	12.56
	307091	AI167439		gb:ox70h06.s1 Soares_NhHMPu_S1 Homo sapi	4.88	8.52
	307181	AI189251		gb:xq99g06.x1 Soares_pregnant_uterus_NbH	3.55	6.44
25	307297	AI205798	Hs.111334	fertilin, light polypeptide	2.46	4.65
	307317	AI208303	Hs.147333	EST	5.64	10.13
	307327	AI214142	Hs.248381	CD68 antigen	3.18	5.15
	307382	AI223158	Hs.147885	ESTs	2.02	3.73
30	307410	AI241715	Hs.77039	ribosomal protein S3A	0.72	0.48
	307415	AI242118		gb:xq92b02.x1 Soares_NFL_T_GBC_S1 Homo s	2.38	3.51
	307423	AI243206	Hs.179573	collagen, type I, alpha 2	2.60	5.44
	307426	AI243364		gb:xq30g11.x1 Soares_NFL_T_GBC_S1 Homo s	3.18	7.67
	307517	AI275055		gb:xq72d03.x1 Soares_NhHMPu_S1 Homo sapi	1.00	1.00
35	307551	AI281556		gb:xq52111.x1 NCL_CGAP_Lym6 Homo sapiens	3.40	11.20
	307561	AI282207		gb:xq65a12.x1 Soares_fetal_lung_NbHL19W	4.74	15.51
	307608	AI290295		gb:xqm01f02.x1 Soares_NhHMPu_S1 Homo sapi	3.50	7.19
	307657	AI306428	Hs.298262	ribosomal protein S19	1.76	2.44
	307691	AI318285		gb:tb17b01.x1 NCL_CGAP_Ov37 Homo sapiens	1.59	1.31
40	307701	AI318583	Hs.276672	EST, Weakly similar to RL6_HUMAN 60S R	1.90	2.13
	307718	AI333405	Hs.83753	small nuclear ribonucleoprotein polypept	0.45	0.99
	307730	AI336092		gb:xq143b07.x1 Soares_fetal_lung_NbHL19W	1.51	0.99
	307760	AI342387		gb:xq127f07.x1 Soares_pregnant_uterus_NbH	1.00	1.00
	307764	AI342731		gb:xq26a07.x1 NCL_CGAP_Lu5 Homo sapiens	4.52	12.58
45	307783	AI347274		gb:ic05d02.x1 NCL_CGAP_Co16 Homo sapiens	1.42	1.00
	307796	AI350556		gb:qt18f09.x1 NCL_CGAP_GC4 Homo sapiens	6.57	9.61
	307807	AI351799		gb:qt09d02.x1 NCL_CGAP_GC4 Homo sapiens	3.38	7.68
	307808	AI351826		gb:qt09g03.x1 NCL_CGAP_GC4 Homo sapiens	0.33	0.86
	307820	AI355761		gb:qt94a11.x1 NCL_CGAP_Co14 Homo sapiens	7.94	21.57
50	307830	AI358722	Hs.276737	EST, Weakly similar to R5HU22 ribosomal	2.05	3.32
	307852	AI365541		gb:xq08g05.x1 NCL_CGAP_CLL1 Homo sapiens	3.18	5.21
	307902	AI380462		gb:tg02h05.x1 NCL_CGAP_CLL1 Homo sapiens	3.13	4.99
	307997	AI434512	Hs.181165	eukaryotic translation elongation factor	1.00	3.01
	308002	AI435240	Hs.283442	ESTs	5.86	12.64
55	308011	AI439473		gb:ti60a08.x1 NCL_CGAP_Lym12 Homo sapien	3.79	5.83
	308023	AI452732	Hs.251577	hemoglobin, alpha 1	0.38	0.88
	308041	AI458824	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.36	6.06
	308059	AI468938	Hs.276877	EST, Weakly similar to RL10_HUMAN 60S R	1.80	1.98
	308085	AI474135	Hs.181165	eukaryotic translation elongation factor	3.38	4.14
60	308101	AI475950	Hs.181165	eukaryotic translation elongation factor	1.30	3.87
	308106	AI476803		gb:ti77e12.x1 Soares_NSF_F8_9W_OT_PA_P_S2.38	8.72	8.72
	308122	AI480123	Hs.309411	EST	2.70	3.86
	308154	AI500600		gb:tn93d08.x1 NCL_CGAP_UI2 Homo sapiens	0.66	1.33
65	308171	AI523632	Hs.298766	ESTs, Weakly similar to schlafen4 [M.mu	2.48	4.86
	308211	AI557029	Hs.278572	anaplastic lymphoma kinase (K1-1)	2.43	2.14
	308213	AI557041		gb:PT2.1_12_E04.r tumor2 Homo sapiens cD	3.34	3.79
	308216	AI557135		gb:PT2.1_13_H06.r tumor2 Homo sapiens cD	4.61	4.78
	308219	AI557246		gb:PT2.1_15_D07.r tumor2 Homo sapiens cD	4.87	7.94
	308271	AI567844	Hs.252259	ribosomal protein S3	2.40	6.35
70	308319	AI583983	Hs.181165	eukaryotic translation elongation factor	2.45	3.33
	308362	AI613519	Hs.105749	KIAA0553 protein	1.24	1.41
	308413	AI636253	Hs.196511	ESTs	3.16	4.82
	308450	AI650860	Hs.96840	KIAA1527 protein	1.79	2.68
	308464	AI672425	Hs.277117	EST, Moderately similar to I38055 myosi	4.87	8.27
75	308588	AI718299		gb:as51g12.x1 Barstead aorta HPLRB6 Homo	3.90	5.64
	308599	AI719893		gb:as47d07.x1 Barstead aorta HPLRB6 Homo	3.32	5.12
	308615	AI738593	Hs.101774	hypothetical protein FLJ23045	3.11	2.36
	308643	AI745040		gb:tr19a12.x1 NCL_CGAP_Ov23 Homo sapiens	3.98	3.69
	308673	AI760864		gb:wi09c10.x1 NCL_CGAP_CLL1 Homo sapiens	0.82	0.99
80	308697	AI767143		gb:wi97a07.x1 NCL_CGAP_Kid12 Homo sapien	2.76	5.59
	308762	AI807405	Hs.259408	ESTs	3.17	6.30
	308778	AI811109		gb:tr04c11.x1 NCL_CGAP_Ov23 Homo sapiens	1.00	1.00
	308782	AI811767	Hs.2186	eukaryotic translation elongation factor	2.94	5.15
	308808	AI818289		gb:wk52c01.x1 NCL_CGAP_Pr22 Homo sapiens	4.41	8.34
	308823	AI824118	Hs.217493	annexin A2	1.85	1.92
85	308875	AI832332		gb:at48g03.x1 Barstead colon HPLRB7 Homo	2.52	3.80

	308879	AI832763	Hs.75968	thymosin, beta 4, X chromosome	3.38	7.96
	308886	AI833240		gb:al76d10.x1 Barstead colon HPLRB7 Homo	3.06	2.65
	308898	AI858845		gb:wl32d10.x1 NCI_CGAP_U11 Homo sapiens	2.45	3.44
5	308934	AI865023	Hs.177	phosphatidylinositol glycan, class H	4.14	6.76
	308966	AI870704		gb:wl47h01.x1 NCI_CGAP_U11 Homo sapiens	1.00	1.00
	308979	AI873111		gb:wl52h05.x1 NCI_CGAP_Bm25 Homo sapien	7.15	11.10
	309045	AI910902		gb:clq39f01.x1 NCI_CGAP_U11 Homo sapiens	0.61	0.59
	309051	AI911975		gb:wd78d01.x1 NCI_CGAP_Lu24 Homo sapiens	1.78	4.42
10	309069	AI917366	Hs.78202	SW/SNF related, matrix associated, act	3.27	5.88
	309083	AI922426	Hs.119598	ribosomal protein L3	2.39	3.34
	309105	AI925503	Hs.265884	ESTs	5.54	17.78
	309122	AI928178		gb:wo95a11.x1 NCI_CGAP_Kd11 Homo sapien	1.00	2.92
	309128	AI928816	Hs.180842	ribosomal protein L13	1.38	5.55
15	309164	AI937761		gb:wp84b09.x1 NCI_CGAP_Bm25 Homo sapien	2.43	3.11
	309177	AI951118		gb:wx63g05.x1 NCI_CGAP_Br18 Homo sapiens	0.81	0.97
	309288	AI991525	Hs.299426	ESTs	4.86	7.46
	309299	AW003478		gb:wwq66c06.x1 NCI_CGAP_GC6 Homo sapiens	4.36	9.43
	309303	AW004823		gb:ws93a08.x1 NCI_CGAP_Co3 Homo sapiens	2.88	7.54
20	309411	AW085201	Hs.244144	EST	4.30	7.14
	309437	AW090702	Hs.278242	tubulin, alpha, ubiquitous	2.49	3.11
	309459	AW117645	Hs.65114	keratin 18	2.88	4.55
	309476	AW129368		gb:xe14b05.x1 NCI_CGAP_U14 Homo sapiens	2.08	6.60
25	309499	AW136325	Hs.279771	Homo sapiens clone PP1596 unknown mRNA	2.82	3.55
	309529	AW150807	Hs.181357	laminin receptor 1 (67kD, ribosomal pro	4.78	3.95
	309532	AW151119		gb:xcg33e10.x1 NCI_CGAP_U11 Homo sapiens	1.18	4.40
	309526	AW192004	Hs.297681	serine (or cysteine) proteinase inhibi	4.46	12.06
	309541	AW194230	Hs.253100	EST, Moderately similar to GHU Ig gamm	1.47	1.39
	309575	AW205681	Hs.253506	EST, Moderately similar to ATPN_HUMAN A	5.68	15.20
30	309593	AW237221	Hs.181357	laminin receptor 1 (67kD, ribosomal prot	1.00	1.00
	309695	AW238011	Hs.295605	mannosidase, alpha, class 2A, member 2	5.45	9.61
	309700	AW241170	Hs.179661	tubulin, beta polypeptide	1.41	1.25
	309747	AW264889		gb:xcq36h02.x1 NCI_CGAP_Lu28 Homo sapiens	5.00	8.35
	309769	AW272346		gb:xc13c10.x1 NCI_CGAP_Kd11 Homo sapien	5.76	11.90
35	309782	AW275156	Hs.156110	immunoglobulin kappa constant	0.42	0.69
	309783	AW275401	Hs.254798	EST	1.00	4.11
	309799	AW276964		gb:xp58h01.x1 NCI_CGAP_Ov39 Homo sapiens	1.68	1.44
	309866	AW299916		gb:xc44c01.x1 NCI_CGAP_Kd11 Homo sapien	3.02	5.04
40	309903	AW339071	Hs.300697	immunoglobulin heavy constant gamma 3 (G	1.05	1.18
	309923	AW340684		gb:hd05g08.x1 Soares_NFL_T_GBC_S1 Homo s	2.30	3.67
	309928	AW341418		gb:hd08c03.x1 Soares_NFL_T_GBC_S1 Homo s	7.41	13.71
	309931	AW341683		gb:hd13d01.x1 Soares_NFL_T_GBC_S1 Homo s	1.20	12.70
	309933	AW341936		gb:hb73f10.x1 NCI_CGAP_U12 Homo sapiens	4.90	18.29
	309964	AW449111	Hs.257111	hypothetical protein MGC3265	1.99	3.07
45	310002	AI439096	Hs.323079	Homo sapiens mRNA: cDNA DKFZp564P116 (fr	0.20	0.47
	310096	AW136822	Hs.172824	ESTs, Weakly similar to B48013 proline-r	1.51	1.22
	310098	AI685841	Hs.161354	ESTs	0.31	0.76
	310109	AI203094	Hs.148633	ESTs	2.06	5.83
	310112	AW197233	Hs.147253	ESTs	2.92	3.55
50	310115	AI611317	Hs.223796	ESTs	1.25	0.84
	310121	AW195642	Hs.148901	ESTs	1.00	2.71
	310146	AI206814	Hs.197422	ESTs	9.50	15.31
	310193	AI627653	Hs.147562	ESTs	2.85	4.18
	310255	AW450439	Hs.153378	ESTs	4.26	10.63
55	310261	AI240483	Hs.201217	ESTs	3.28	4.40
	310264	AI915771	Hs.74170	metallothionein 1E (functional)	0.26	0.86
	310275	AI242102	Hs.213636	ESTs	5.43	8.19
	310282	AI243332	Hs.156055	ESTs	3.15	8.06
	310290	AW013815	Hs.149103	ESTs	2.19	3.12
60	310333	AI253200	Hs.145402	ESTs	1.17	1.91
	310346	AI261340	Hs.145517	ESTs	4.81	9.95
	310385	AI263392	Hs.156151	ESTs	5.96	7.79
	310443	AW119018	Hs.164231	ESTs	2.90	4.63
	310444	AW196632	Hs.252956	ESTs	0.85	1.01
65	310446	AI275715	Hs.145926	ESTs	2.18	3.85
	310468	AI984074	Hs.196398	ESTs	3.39	5.19
	310477	AI948801	Hs.171073	ESTs	1.00	1.00
	310512	AW275603	Hs.200712	ESTs	3.87	8.12
	310514	AI681145	Hs.160724	ESTs	3.30	7.33
70	310524	AW082270	Hs.12496	ESTs, Highly similar to AC004836 1 simil	0.72	1.44
	310547	AI302654	Hs.208024	ESTs	3.26	3.46
	310584	AI653007	Hs.156304	ESTs	2.39	4.08
	310608	AI962234	Hs.196102	ESTs	5.60	6.49
	310624	AI341594		gb:Human endogenous retrovirus H proteas	4.91	9.09
75	310636	AI814373	Hs.154175	ESTs	1.85	1.71
	310648	AI347853	Hs.156672	ESTs	0.17	0.69
	310694	AI654370	Hs.157752	Homo sapiens mRNA full length insert cDN	5.40	13.22
	310695	AI472124	Hs.157757	ESTs	4.82	6.27
	310714	AI418446	Hs.157882	ESTs	1.76	3.51
	310722	AI989803	Hs.157289	ESTs	1.14	6.85
80	310756	AI916560	Hs.158707	ESTs	8.46	13.01
	310764	AI376769	Hs.167172	ESTs	4.76	7.37
	310848	AI459554	Hs.161286	ESTs	2.84	1.96
	310851	AW291714	Hs.221703	ESTs	1.00	2.32
85	310854	AI421677	Hs.161332	ESTs	6.37	7.94
	310858	AI871000	Hs.161330	ESTs	6.07	9.84

	310864	AI924558	Hs.161399	ESTs	0.87	0.78
	310875	T47764	Hs.132917	ESTs	1.00	3.63
	310896	AW157731	Hs.270982	ESTs, Moderately similar to ALU7_HUMAN A	7.07	16.68
5	310922	AW195634	Hs.170401	ESTs	1.00	1.00
	310955	AI560210	Hs.263912	ESTs	10.08	17.66
	310957	AW190974	Hs.196918	ESTs	2.18	3.18
	311000	AI521830	Hs.171050	ESTs	3.06	6.64
	311012	AW299070	Hs.241097	ESTs	1.23	3.77
10	311034	AI564023	Hs.311389	ESTs, Moderately similar to PT0375 natur	2.44	2.09
	311074	AW290922	Hs.199848	ESTs	6.04	14.19
	311134	AI990849	Hs.196971	ESTs	3.54	6.96
	311174	AW450552	Hs.205457	peritaxin	0.65	0.95
	311187	AI638374	Hs.224189	ESTs	2.46	2.78
15	311220	AI656040	Hs.196532	ESTs	1.10	2.52
	311230	AI989808	Hs.197663	ESTs	1.41	1.75
	311236	AI653378	Hs.197674	ESTs	2.18	2.11
	311242	AW016812	Hs.200266	ESTs	0.63	5.11
	311258	AI671221	Hs.199887	ESTs	1.00	1.41
20	311277	AW072813	Hs.270868	ESTs, Moderately similar to ALU4_HUMAN A	2.56	1.94
	311294	AA826425	Hs.291829	ESTs	1.04	2.69
	311308	F12664	Hs.49000	ESTs	1.96	6.70
	311351	AI682303	Hs.201274	ESTs	4.77	9.38
	311390	AW392997	Hs.202280	ESTs	2.80	6.06
25	311405	AW290961	Hs.201815	ESTs	3.80	11.66
	311409	AI698839		gb:wd31f02.x1 Soares_NFL_T_GBC_S1 Homo s	3.84	6.94
	311420	AI936291	Hs.209867	ESTs	5.30	12.56
	311443	AI791521	Hs.192206	ESTs	4.39	6.09
	311467	AI934909	Hs.175377	ESTs	1.00	1.04
30	311479	AI933672	Hs.211399	ESTs	2.76	5.61
	311488	R57390	Hs.301064	arfaptin 1	2.50	5.73
	311495	AW300077	Hs.221358	ESTs	3.63	6.09
	311511	AW444568	Hs.210303	ESTs	2.00	2.87
	311534	AW130351	Hs.243549	ESTs	0.31	1.33
35	311537	AI805121	Hs.211828	ESTs	3.69	5.85
	311543	AI681360	Hs.201259	ESTs	1.73	1.34
	311551	AW449774	Hs.296380	POM (POM121 rat homolog) and ZP3 fusion	3.31	6.12
	311557	AI819230	Hs.211238	interleukin-1 homolog 1	1.00	1.00
	311558	Z44432	Hs.63128	KIAA1292 protein	2.25	3.41
40	311559	AW008271	Hs.265848	similar to rat myomegalin	2.68	5.90
	311563	AI922143	Hs.211334	ESTs	2.39	3.32
	311586	AI827834	Hs.211227	ESTs	2.47	3.85
	311616	AW450675	Hs.212709	ESTs	1.00	1.00
	311621	AI924307	Hs.213464	ESTs	4.16	6.74
45	311635	AI928456	Hs.213081	ESTs	2.17	3.76
	311668	AW193674	Hs.240044	ESTs	2.60	3.12
	311672	R11807	Hs.20914	hypothetical protein FLJ23056	2.79	5.18
	311683	AW183738	Hs.232644	ESTs	0.19	0.96
	311700	R49601	Hs.171495	retinoic acid receptor, beta	6.28	8.83
50	311714	AW131785	Hs.246831	ESTs, Weakly similar to CLK_HUMAN VOLTA	5.00	8.17
	311735	AW294416	Hs.144687	Homo sapiens cDNA FLJ12981 fis, clone NT	0.96	0.72
	311743	T99079	Hs.191194	ESTs	1.00	1.95
	311783	AI682478	Hs.13528	hypothetical protein FLJ14054	0.16	0.77
	311785	AI056769	Hs.133512	ESTs	1.34	3.97
55	311799	AA780791	Hs.14014	ESTs, Weakly similar to KIAA0973 protein	8.52	13.32
	311819	AW265275	Hs.254325	ESTs	3.58	3.91
	311823	AI089422	Hs.131297	ESTs	1.40	1.72
	311877	AA349893	Hs.85339	G protein-coupled receptor 39	0.95	0.91
	311886	AA522738	Hs.132554	ESTs	0.88	0.87
60	311896	AW206447		gb:U1-H-B11-atg-g-02-0-U1.s1 NCL CGAP_Su	1.66	1.13
	311910	N28365	Hs.22579	Homo sapiens clone CDABP0036 mRNA sequen	1.66	2.30
	311923	T60843	Hs.189679	ESTs	0.42	2.63
	311933	AI597963	Hs.118726	ESTs	1.88	3.02
	311959	T67262	Hs.124733	ESTs	2.02	2.33
65	311960	AW440133	Hs.189690	ESTs	3.87	6.62
	311967	AI392726	Hs.182434	ESTs	5.80	8.14
	311975	AA804374	Hs.272203	Homo sapiens cDNA FLJ20843 fis, clone AD	0.98	3.26
	312005	T78450	Hs.13941	ESTs	0.12	1.39
	312028	T78886	Hs.284450	ESTs	3.78	4.92
70	312046	AI580018	Hs.268591	ESTs	4.11	7.32
	312056	T83748	Hs.268594	ESTs	2.36	3.08
	312064	AA676713	Hs.191155	ESTs	3.34	5.28
	312088	AW303760	Hs.13685	ESTs	1.60	1.15
	312093	T91809	Hs.121296	ESTs	0.68	0.85
75	312094	Z78390		gb:HSZ78390 Human fetal brain S. Meier-E	3.05	4.48
	312097	AI352096	Hs.112180	zinc finger protein 148 (pH2-52)	4.52	9.70
	312118	T85332	Hs.178294	ESTs	2.40	2.60
	312128	AI052609	Hs.17631	Homo sapiens cDNA FLJ20118 fis, clone CO	2.39	3.53
	312147	T89855	Hs.195648	ESTs	0.67	1.03
80	312175	AA953383	Hs.127554	ESTs	5.85	10.60
	312179	AI052572	Hs.269864	ESTs	2.41	3.32
	312201	AI928365	Hs.91139	solute carrier family 1 (neuronal/epithel	0.24	0.89
	312207	H90213	Hs.191330	ESTs	2.20	4.55
	312220	N74513		gb:za55a07.s1 Soares fetal liver spleen	4.28	11.13
85	312252	AI126388	Hs.143655	ESTs	1.64	1.57
	312304	AA491949	Hs.269392	ESTs	0.12	2.47

	312318	AW235092	Hs.143981	ESTs	3.45	5.69
	312319	AA216698	Hs.180780	TERA protein	5.78	4.46
	312321	R65210	Hs.186937	ESTs	0.44	1.74
5	312331	AA825512	Hs.289101	glucose regulated protein, 58kD	3.73	5.96
	312339	AA524394	Hs.165544	ESTs	3.07	0.95
	312353	AI675558	Hs.181857	ESTs	10.08	16.73
	312375	AI375095	Hs.172405	cell division cycle 27	2.78	3.71
	312376	R52089	Hs.172717	ESTs	1.00	1.00
10	312389	AI863140		gbt243h12x1 NCL CGAP_Bm52 Homo sapien	2.37	3.98
	312437	AA995028		gb:RC4-BT0629-120200-011-b10 BT0629 Homo	4.06	5.41
	312440	AI051133	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f	1.00	1.00
	312451	R59989	Hs.176539	ESTs	4.96	10.04
	312458	AI167637	Hs.146924	ESTs	1.11	1.00
	312507	AI168177	Hs.143653	ESTs	5.89	8.24
15	312520	AI742591	Hs.205392	ESTs	3.30	8.92
	312548	AI566228	Hs.159426	hypothetical protein PRO2121	1.38	1.65
	312564	H21520	Hs.35088	ESTs	0.40	0.77
	312583	AI193122	Hs.124141	ESTs	0.13	0.94
20	312599	AI865073	Hs.125720	ESTs	3.75	5.29
	312602	AA046451	Hs.165200	ESTs	6.78	12.93
	312645	H52121	Hs.193007	ESTs	0.38	1.13
	312666	AI240582	Hs.214678	ESTs	0.98	2.03
	312689	AW450461	Hs.203665	ESTs	0.21	0.61
	312817	H75459	Hs.233425	ESTs	1.51	0.85
25	312846	AW152104	Hs.200879	ESTs	8.93	13.78
	312873	AI690071	Hs.283552	ESTs, Weakly similar to unnamed protein	4.20	6.23
	312893	AI016204	Hs.172922	ESTs	2.67	3.15
	312902	AW292797	Hs.130316	ESTs, Weakly similar to T2D3_HUMAN TRANS	1.19	0.71
30	312925	N90868	Hs.271695	ESTs	2.50	4.25
	312936	AI681581	Hs.121525	ESTs	1.00	1.17
	312975	AI640506	Hs.293119	ESTs, Weakly similar to ALU7_HUMAN ALU S	2.30	4.80
	312978	N24887	Hs.292500	ESTs	0.80	1.05
	312980	AA497043	Hs.115685	ESTs	3.12	3.60
35	312984	N25871	Hs.177337	ESTs	2.03	2.13
	313000	AI147412	Hs.146657	ESTs	5.52	8.42
	313029	AA731520	Hs.170504	ESTs	0.96	1.39
	313039	AI419290	Hs.149990	ESTs, Weakly similar to unnamed protein	6.48	13.20
	313049	AW293055	Hs.119357	ESTs	6.44	10.73
40	313056	AI651930	Hs.135584	ESTs	1.51	2.04
	313058	D81015	Hs.125382	ESTs	0.25	1.50
	313070	AI422023	Hs.161338	ESTs	8.55	11.60
	313097	AI676164	Hs.204339	ESTs	3.72	4.56
	313130	AW449171	Hs.168677	ESTs	3.28	5.06
45	313136	N59284	Hs.288010	ESTs	0.49	1.36
	313153	AI240838	Hs.132750	ESTs	5.36	5.52
	313210	N74077	Hs.197043	ESTs	0.30	0.66
	313236	AW238169	Hs.83513	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.16	8.76
	313239	W19632	Hs.124170	ESTs	1.00	3.87
50	313265	N93466	Hs.121764	ESTs, Weakly similar to testicular tektd	0.74	2.06
	313267	AI770008	Hs.129583	ESTs	0.23	1.30
	313275	AI027604	Hs.159650	ESTs	6.68	9.57
	313290	AI753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	1.34	1.07
	313292	AI362991	Hs.202121	ESTs, Weakly similar to env protein [Hs	2.00	4.32
55	313325	AI420611	Hs.127832	ESTs	1.20	2.27
	313357	AW074848	Hs.201501	ESTs	4.02	5.33
	313393	AI674685	Hs.200141	ESTs	1.36	2.84
	313399	AW376889	Hs.194097	ESTs	2.58	5.26
	313414	AI241540	Hs.132933	ESTs	6.57	15.07
60	313417	AA741151	Hs.137323	ESTs	0.63	3.01
	313457	AA576052	Hs.193223	Homo sapiens cDNA FLJ11646 fis, clone HE	2.78	4.70
	313499	AI261390	Hs.146085	KIAA1345 protein	0.91	2.37
	313516	AA029058	Hs.135145	ESTs	3.41	7.08
	313556	AA628517	Hs.118502	ESTs	0.23	0.70
65	313569	AI273419	Hs.135146	hypothetical protein FLJ13984	1.89	1.00
	313570	AA041455	Hs.209312	ESTs	0.73	2.27
	313638	AI753075	Hs.104627	Homo sapiens cDNA FLJ10158 fis, clone HE	1.00	1.72
	313662	AA740151	Hs.130425	ESTs	0.20	1.42
	313671	W49823	Hs.104613	RP42 homolog	1.00	1.00
70	313672	AW468891	Hs.122948	ESTs	3.46	5.80
	313690	AI493591	Hs.78146	platelet/endothelial cell adhesion molec	0.51	0.97
	313711	AA398070	Hs.133471	ESTs	0.18	1.01
	313723	AA070412		gb:zm68c10.s1 Stratagene neuroepithelium	1.08	1.03
	313726	AI744687	Hs.257806	ESTs	2.13	2.99
75	313774	AW136836	Hs.144583	ESTs	1.38	1.19
	313784	AA910514	Hs.134905	ESTs	3.88	5.78
	313790	AW078569	Hs.177043	ESTs	0.22	2.06
	313832	AW271022	Hs.133294	ESTs	1.15	0.91
	313834	AW418779	Hs.114889	ESTs	0.68	3.14
80	313835	AI538438	Hs.159087	ESTs	5.74	8.88
	313852	H18633	Hs.123641	protein tyrosine phosphatase, receptor t	0.16	1.14
	313854	AW470806	Hs.275002	ESTs	2.09	4.06
	313865	AA731470	Hs.163839	ESTs	3.41	4.09
	313871	AW471088	Hs.145950	ESTs	5.28	6.83
85	313883	AI949384		gb:nu76d01.s1 NCL CGAP_AIv1 Homo sapiens	2.90	10.91
	313915	AI969390	Hs.163443	Homo sapiens cDNA FLJ11576 fis, clone HE	1.00	1.00

	313926	AW473830	Hs.171442	ESTs	3.40	4.11
	313948	AW452823	Hs.135268	ESTs	5.77	9.15
	313978	AJ870175	Hs.13957	ESTs	0.46	0.75
5	313983	AJ829133	Hs.226780	ESTs	4.10	6.40
	314035	AA164199	Hs.270152	ESTs	5.88	7.90
	314037	AW300048	Hs.275272	ESTs	1.00	3.79
	314040	AA166970	Hs.118748	ESTs	7.60	11.33
	314067	AW293538	Hs.51743	KIAA1340 protein	1.85	1.21
10	314103	AJ028477	Hs.132775	ESTs	2.90	5.29
	314107	AA806113	Hs.189025	ESTs	2.00	1.66
	314113	AA218986	Hs.118854	ESTs	0.91	4.17
	314124	AW118745	Hs.9460	Homo sapiens mRNA; cDNA DKFZp547C244 (fr	2.53	3.32
	314126	AA226431	gb:nc16b12.s1 NCL_CGAP_Pr1 Homo sapiens	3.13	5.08	
15	314128	AA935633	Hs.194628	ESTs	2.90	6.35
	314151	AA236163	Hs.202430	ESTs	4.15	6.45
	314184	AW081795	Hs.233465	ESTs	3.44	4.65
	314192	AW290975	Hs.116923	ESTs	1.00	1.23
	314244	AL036450	Hs.103238	ESTs	2.88	3.67
20	314253	AA278679	Hs.189510	ESTs	4.98	7.16
	314262	AW086215	Hs.246096	ESTs	0.38	1.94
	314320	AA811598	Hs.275809	ESTs	3.34	5.66
	314332	AL037551	Hs.95612	ESTs	2.85	2.09
	314335	AA287443	Hs.142570	Homo sapiens clone 24629 mRNA sequence	4.35	4.78
25	314340	AW304350	Hs.130879	ESTs, Moderately similar to putative p15	0.77	0.86
	314351	AA292275	Hs.193746	ESTs	3.07	3.77
	314376	AJ626633	Hs.324679	ESTs	4.10	6.11
	314443	AA827125	Hs.192043	ESTs	6.20	13.67
	314458	AI217440	Hs.143873	ESTs	0.58	2.49
30	314456	AA767818	Hs.122707	ESTs	2.53	2.62
	314478	AI521173	Hs.125507	DEAD-box protein	3.94	5.65
	314482	AL043807	Hs.134182	ESTs	1.30	1.44
	314506	AA833655	Hs.206868	Homo sapiens cDNA FLJ14056 fis, clone HE	3.28	3.47
	314519	RA2554	Hs.210862	T-box, brain, 1	3.12	6.16
35	314529	AL046412	Hs.202151	ESTs	3.43	6.87
	314546	AW007211	Hs.16131	hypothetical protein FLJ12876	1.38	1.00
	314562	AI584127	Hs.143493	ESTs	2.29	5.27
	314579	AW197442	Hs.116958	ESTs	3.87	5.75
	314580	AW451832	Hs.265938	ESTs, Moderately similar to KIAA1200 pro	0.10	0.71
40	314585	AA918474	Hs.216363	ESTs	1.08	1.40
	314589	AW384790	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	1.00	1.00
	314592	AA435761	Hs.192148	ESTs	0.90	2.60
	314603	AA418024	Hs.270670	ESTs	4.56	6.29
	314604	AA946582	Hs.8700	deleted in liver cancer 1	3.42	3.92
45	314606	AA418241	Hs.188767	ESTs	2.97	4.55
	314648	AA878419		gb:EST391378 MAGE resequences, MAGP Homo1.42	1.36	1.36
	314699	AJ038719	Hs.132801	ESTs	3.66	4.97
	314701	AJ754634	Hs.131987	ESTs	0.03	0.90
	314710	AJ669131	Hs.290989	EST	3.40	7.52
50	314750	AI095005	Hs.135174	ESTs	2.80	6.54
	314767	AW135412	Hs.164002	ESTs	3.20	4.26
	314801	AA481027	Hs.109045	hypothetical protein FLJ10498	1.00	1.00
	314817	AJ694139	Hs.192855	ESTs	0.91	0.99
	314835	AJ281370	Hs.76064	ribosomal protein L27a	5.75	7.44
55	314852	AI903735		gb:MR-BT035-200199-031 BT035 Homo sapien	1.68	4.34
	314853	AA729232	Hs.153279	ESTs	0.60	1.85
	314940	AW452768	Hs.162045	ESTs	10.10	16.20
	314941	AA515802	Hs.130650	ESTs	0.31	1.02
	314943	AJ476797	Hs.184572	cell division cycle 2, G1 to S and G2 to	2.18	0.37
60	314955	AA521382	Hs.192534	ESTs	2.59	3.90
	314973	AW273128	Hs.300268	ESTs	1.05	1.25
	315004	AA527941	Hs.325351	EST	5.64	13.63
	315006	AI538613	Hs.298241	Transmembrane protease, serine 3	0.52	1.78
	315033	AI493046	Hs.146133	ESTs	2.46	1.00
65	315035	AI569476	Hs.177135	ESTs	0.34	1.33
	315056	AI202703	Hs.152414	ESTs	2.10	2.64
	315069	AI821517	Hs.105866	ESTs	1.00	1.30
	315071	AA552690	Hs.152423	Homo sapiens cDNA: FLJ21274 fis, clone C	1.78	1.00
	315073	AW452948	Hs.257631	ESTs	1.17	1.52
70	315078	AA568548	Hs.190616	ESTs	3.00	3.79
	315080	AA744550	Hs.136345	ESTs	1.00	1.00
	315120	AA564991	Hs.269477	ESTs	0.64	1.44
	315175	AI025842	Hs.152530	ESTs	0.61	1.91
	315193	AI241331	Hs.131765	ESTs	1.06	0.97
75	315196	AA972756	Hs.44898	Homo sapiens clone TCCCTA00151 mRNA sequ	0.48	1.96
	315200	AI808235	Hs.307686	EST	3.76	9.40
	315254	AI474433	Hs.179556	ESTs	5.37	9.36
	315353	AW452608	Hs.279610	hypothetical protein FLJ10493	1.00	1.30
	315397	AA218940	Hs.137516	fidgetin-like 1	3.38	2.24
80	315403	AW362980	Hs.163924	ESTs	2.04	5.23
	315431	AA622104	Hs.184838	ESTs	2.36	8.04
	315454	AI239473		gb:zh36f02.x1 Soares_NFL_T_GBC_S1 Homo s	3.46	7.64
	315455	AW393391	Hs.156919	ESTs	3.78	5.76
	315473	AI681671	Hs.312671	ESTs, Moderately similar to OVCA1	0.89	2.15
	315483	AW512763	Hs.222024	transcription factor BMAL2	2.32	1.96
85	315526	AI193048	Hs.128585	ESTs	1.67	1.78

	315530	AI200852	Hs.127780	ESTs	1.05	1.01
	315541	AI168233	Hs.123159	sperm associated antigen 4	0.85	0.56
	315552	AW445034	Hs.256578	ESTs	1.00	2.22
5	315562	AA737415	Hs.152826	ESTs	2.66	2.48
	315577	AW513545	Hs.17283	hypothetical protein FLJ10690	2.20	2.25
	315587	AI268399	Hs.140489	ESTs	1.00	1.04
	315589	AW072387	Hs.158258	Homo sapiens mRNA; cDNA DKFZp434B1272 (f	0.14	1.05
	315623	AA364078	Hs.258189	ESTs	7.44	12.56
10	315634	AA837085	Hs.220585	ESTs	0.50	1.40
	315658	AA912347	Hs.136585	ESTs	0.43	1.22
	315677	AI932662	Hs.164073	ESTs	0.60	1.39
	315706	AW440742	Hs.155556	hypothetical protein FLJ20202	2.18	3.77
	315707	AI418055	Hs.161160	ESTs	2.88	2.63
15	315730	H25899	Hs.201591	ESTs	0.11	0.60
	315745	AI821759	Hs.191856	ESTs	3.50	7.25
	315791	AA678177	gb:U15a05.s1 Soares_fetal_liver_spleen_	1.78	2.63	
	315801	AA827752	Hs.266134	ESTs	4.31	6.23
	315820	AI652022	Hs.258785	ESTs	2.35	3.01
20	315878	AA683336	Hs.189046	ESTs	2.12	2.64
	315905	AI821911	Hs.209452	ESTs	1.03	1.97
	315923	AI052789	Hs.133263	ESTs	2.63	5.05
	315954	AW276810	Hs.254859	ESTs, Moderately similar to ALU5_HUMAN A	1.21	0.85
	315978	AA830893	Hs.119769	ESTs	3.09	3.41
25	316001	AI248584	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone C	2.20	6.82
	316011	AW516953	Hs.201372	ESTs	0.35	1.63
	316012	AA764950	Hs.119898	ESTs	6.56	8.13
	316040	AI983409	Hs.189226	ESTs	5.69	10.69
	316048	AI720759	Hs.224971	ESTs	2.84	10.45
30	316076	AW297895	Hs.116424	ESTs	0.30	1.05
	316124	AI308862	Hs.167029	ESTs	1.00	1.43
	316151	AI806016	Hs.156520	ESTs	5.80	9.03
	316187	AW518299	Hs.192253	ESTs	1.20	3.96
	316204	AA731509	Hs.120257	ESTs	4.92	6.94
35	316232	AW297853	Hs.251203	ESTs	1.48	1.60
	316275	AI671041	Hs.292611	ESTs, Moderately similar to ALU1_HUMAN A	5.86	12.14
	316291	AW375974	Hs.156704	ESTs	2.73	2.69
	316303	AA740994	Hs.209609	ESTs	1.53	1.26
	316344	AA744518	Hs.120610	ESTs	3.66	8.34
40	316346	AI028478	Hs.157447	ESTs	3.51	6.69
	316365	AI627845	Hs.210776	ESTs	2.50	4.33
	316380	AI393378	Hs.164496	ESTs	1.16	2.16
	316470	AA809902	Hs.243813	ESTs	5.40	10.34
	316509	AA767310	Hs.291766	ESTs	2.46	2.89
45	316514	AA768037	Hs.291671	ESTs	4.70	6.04
	316519	AI929097	gb:cd10c11.s1 NCI_CGAP_GC81 Homo sapiens	4.41	9.70	
	316609	AW292520	Hs.122082	ESTs	1.00	2.89
	316633	AI125586	Hs.127955	ESTs	2.61	3.72
	316700	AW172316	Hs.252961	ESTs, Weakly similar to ALU1_HUMAN ALU S	3.46	4.64
50	316711	AI743721	Hs.285316	ESTs, Moderately similar to ALU7_HUMAN A	4.45	6.95
	316713	AI090671	Hs.134807	hypothetical protein FLJ12057	0.30	2.40
	316715	AI440266	Hs.170673	ESTs, Weakly similar to AF126780 1 refn	0.20	1.45
	316787	AW369770	Hs.130351	ESTs	4.05	5.53
	316809	AA825839	Hs.202238	ESTs	2.25	3.82
55	316811	AA922060	Hs.132471	ESTs	1.00	1.32
	316812	AW135045	Hs.232001	ESTs	3.28	4.70
	316818	AA827176	Hs.124316	ESTs	0.67	1.81
	316824	AA837416	Hs.124299	ESTs	3.53	6.00
	316827	AI380429	Hs.172445	ESTs	0.72	1.56
60	316891	AW298119	Hs.202536	ESTs	1.64	2.97
	316951	AA134365	Hs.57548	ESTs	1.45	1.08
	316970	AA860172	Hs.132406	ESTs	1.00	1.53
	316971	AA860212	Hs.170991	ESTs	1.08	1.96
	316990	AA861611	Hs.130643	ESTs	5.44	10.04
65	317001	AI627917	Hs.233694	hypothetical protein FLJ11350	3.56	4.37
	317008	AW051597	Hs.143707	ESTs	0.69	1.37
	317051	AA873253	Hs.126233	ESTs	6.18	12.72
	317128	AA971374	Hs.125674	ESTs	1.87	2.66
	317129	HI12523	Hs.78521	Homo sapiens cDNA: FLJ21193 fis, clone C	4.12	6.64
70	317137	AW341567	Hs.125710	ESTs	2.82	5.12
	317196	AI348258	Hs.153412	ESTs	1.98	2.51
	317212	AI866468	Hs.148294	ESTs	1.86	2.83
	317223	AW297920	Hs.130054	ESTs	0.83	1.57
	317224	D56760	Hs.93029	sparc:osteoneclin, cwcv and kazal-like d	2.74	0.86
75	317266	AA906289	Hs.203614	ESTs	1.00	1.00
	317282	AI807444	Hs.176101	ESTs	2.60	4.21
	317285	AW370882	Hs.222080	ESTs	1.96	3.49
	317302	AA908709	Hs.135564	ESTs	7.16	8.32
	317304	AW449899	Hs.130184	ESTs	1.38	2.28
80	317320	AA927151	Hs.130452	ESTs	3.58	8.13
	317413	AW341701	Hs.126622	ESTs	2.08	4.92
	317417	AA918420	Hs.145378	ESTs	3.06	4.79
	317452	AA972955	Hs.135568	ESTs	4.22	9.21
	317519	AI859695	Hs.126860	ESTs	1.88	4.15
85	317521	AI824338	Hs.126891	ESTs	3.12	4.55
	317529	AI916517	Hs.126865	ESTs	2.73	3.34

	317570	AI733361	Hs.127122	ESTs	1.00	2.43
	317571	AA938663	Hs.199828	ESTs	5.20	11.95
	317598	AW206035	Hs.192123	ESTs	0.33	1.56
5	317627	AI346110	Hs.132553	ESTs	1.50	1.39
	317650	AI733310	Hs.127346	ESTs	0.48	1.46
	317659	AA961216	Hs.127785	ESTs	4.18	7.14
	317674	AW294909	Hs.132208	ESTs	2.92	3.20
	317686	AA969051	Hs.187319	ESTs	1.00	1.01
10	317692	AI307659	Hs.174794	ESTs	5.33	9.59
	317701	AI674774	Hs.128014	ESTs	1.00	1.00
	317711	AI733015	Hs.272189	ESTs	5.13	7.81
	317722	AI733373	Hs.128119	ESTs	2.50	6.03
	317756	AA973667	Hs.128320	ESTs	1.59	1.30
15	317777	AI143525	Hs.47313	KIAA0258 gene product	1.00	2.48
	317799	AA982773	Hs.128808	ESTs	1.78	2.11
	317803	AA983251	Hs.128899	ESTs	0.80	1.06
	317821	AI368158	Hs.70983	PTPL1-associated RhoGAP 1	0.17	0.68
	317848	AI820575	Hs.129088	Homo sapiens cDNA FLJ12007 fis, clone HE	5.30	8.16
20	317850	N29974	Hs.152982	hypothetical protein FLJ13117	1.30	2.28
	317861	AW341064	Hs.129119	ESTs	2.18	5.93
	317865	AI298794	Hs.129130	ESTs	4.48	8.20
	317869	AW295184	Hs.129142	deoxyribonuclease II beta	0.44	0.99
	317881	AI827248	Hs.224398	Homo sapiens cDNA FLJ11469 fis, clone HE	4.06	2.23
25	317890	AI915599	Hs.129225	ESTs	4.68	7.48
	317899	AI952430	Hs.150614	ESTs, Weakly similar to ALU4_HUMAN ALU S	3.14	3.37
	317986	AI005163	Hs.201378	ESTs, Weakly similar to T12545 hypotheti	0.28	1.66
	318001	AW235697	Hs.130980	ESTs	5.12	9.97
	318016	AI016694	Hs.256921	ESTs	1.86	4.50
	318023	AW243058	Hs.131155	ESTs	2.92	5.22
30	318054	AW449270	Hs.232140	ESTs	3.92	6.37
	318058	AI024540	Hs.131574	ESTs	1.21	1.27
	318117	AI208304	Hs.250114	ESTs	0.86	1.17
	318187	AI792585	Hs.133272	ESTs, Weakly similar to ALUC_HUMAN IIII	5.90	6.98
35	318223	AI077540	Hs.134090	ESTs	1.05	0.90
	318240	AI085377	Hs.143610	ESTs	3.10	2.40
	318255	AI082692	Hs.134652	ESTs	0.02	1.05
	318266	AI554341	Hs.271443	ESTs	6.12	10.55
	318330	AI093840	Hs.143758	ESTs	4.98	7.90
40	318369	AI493501	Hs.170974	ESTs	2.46	5.62
	318428	AI949409	Hs.194591	ESTs	0.77	0.45
	318458	AI149783	Hs.158438	ESTs	3.54	4.92
	318467	AI151395	Hs.144834	ESTs	4.56	5.62
	318473	AI939339	Hs.146883	ESTs	2.08	4.05
45	318476	AI693927	Hs.265165	ESTs	4.22	8.07
	318487	AI167877	Hs.143716	ESTs	1.47	1.05
	318488	AI217431	Hs.144709	ESTs	1.40	4.14
	318491	T26477	Hs.22883	ESTs, Weakly similar to ALU8_HUMAN ALU S	1.84	1.90
	318499	T25451		gb:PTHI188 HTCCL1 Homo sapiens cDNA 5/3	2.58	5.20
50	318537	AA377908	Hs.13254	ESTs	3.26	4.18
	318538	N28625	Hs.74034	Homo sapiens clone 24651 mRNA sequence	0.35	1.07
	318547	R20578	Hs.90431	ESTs	3.22	4.60
	318552	R18364	Hs.90363	ESTs	4.87	9.05
	318575	R55102	Hs.107761	ESTs, Weakly similar to unnamed protein	1.91	1.98
55	318580	T34571	Hs.49007	poly(A) polymerase alpha	2.74	6.22
	318587	AA779704	Hs.168830	Homo sapiens cDNA FLJ12136 fis, clone MA	0.85	2.46
	318596	AI470235	Hs.172698	EST	4.88	4.93
	318622	T48325	Hs.237658	apolipoprotein A-II	4.80	12.51
	318629	N25163	Hs.8861	ESTs	0.39	1.04
60	318637	AA243539	Hs.9196	hypothetical protein	1.72	3.57
	318648	T77141	Hs.184411	albumin	6.27	9.91
	318650	AA393302	Hs.176626	hypothetical protein EDAG-1	3.96	8.84
	318671	AA188823	Hs.299254	Homo sapiens cDNA: FLJ23597 fis, clone L	1.53	0.81
	318679	T58115	Hs.10336	ESTs	1.00	2.19
65	318711	AI936475	Hs.101282	Homo sapiens cDNA: FLJ21238 fis, clone C	3.05	3.18
	318725	AI962487	Hs.242990	ESTs	1.08	2.46
	318728	Z30201	Hs.291289	ESTs, Weakly similar to ALU1_HUMAN ALU S	0.77	1.33
	318740	NM_002543	Hs.77729	oxidised low density lipoprotein (lectin	0.25	1.49
	318776	R24963	Hs.23766	ESTs	1.00	3.01
70	318784	H00148	Hs.5181	proliferation-associated 2G4, 38kD	2.70	3.86
	318816	F07873	Hs.21273	ESTs	3.90	7.13
	318865	H10818		gb:ym04f10.r1 Soares Infant brain 1N1B H	2.25	3.56
	318879	R56332	Hs.18268	adenylate kinase 5	1.78	5.00
	318881	Z43224	Hs.124952	ESTs	4.79	14.13
75	318894	F08138	Hs.7387	DKFZP564B116 protein	5.31	7.00
	318901	AW368520	Hs.301528	L-kynurenine/alpha-aminoadipate aminotra	1.03	0.91
	318925	Z43577	Hs.21470	ESTs	2.23	3.80
	318936	AI219221	Hs.308298	ESTs	1.86	7.16
	318982	Z44140	Hs.269622	ESTs	5.84	9.79
80	318986	Z44186	Hs.169161	ESTs, Highly similar to MAON_HUMAN NADP-	1.00	1.00
	319041	Z44720	Hs.98365	ESTs, Weakly similar to weak similarity	3.38	6.11
	319103	H05856	Hs.4993	KIAA1313 protein	1.00	1.07
	319170	R13678	Hs.285306	putative selenocysteine lyase	3.79	5.03
	319196	F07953	Hs.16085	putative G-protein coupled receptor	1.00	2.98
85	319199	F07361	Hs.13306	ESTs	3.53	5.66
	319242	F11472	Hs.12839	ESTs	5.87	7.26

	319263	T65331	Hs.81360	Homo sapiens cDNA: FLJ21927 fis, clone H	1.81	1.57
	319267	F11802	Hs.6818	ESTs	1.10	4.72
	319270	R13474	Hs.290263	ESTs	4.80	10.40
5	319279	T65094	Hs.12677	CGI-147 protein	1.50	2.11
	319282	AA461358	Hs.12876	ESTs	1.00	1.00
	319289	W07304	Hs.79059	transforming growth factor, beta recepto	0.18	0.68
	319291	W86578	Hs.285243	hypothetical protein FLJ22029	0.26	0.62
	319293	F12119	Hs.12583	ESTs	3.13	4.50
10	319312	Z45481		gb:HSC2QE041 normalized infant brain cDN	1.10	1.00
	319370	H54254	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	0.16	0.73
	319391	R06304	Hs.13911	ESTs	1.26	2.43
	319396	H67130	Hs.301743	ESTs	0.70	0.76
	319398	AA359754	Hs.191196	ESTs	2.45	3.59
15	319407	R05329		gb:ye91b04.r1 Soares fetal liver spleen	2.00	3.54
	319425	T82930		gb:yd39i07.r1 Soares fetal liver spleen	4.28	8.81
	319433	R06050	Hs.191198	ESTs	6.15	14.13
	319437	AA282420	Hs.111991	ESTs, Weakly similar to Y48A5A.1 (C.eleg	3.26	5.68
	319466	A809937	Hs.116417	ESTs	1.76	5.65
20	319471	R06546	Hs.19717	ESTs	4.29	4.84
	319480	R06933	Hs.184221	ESTs	1.00	1.00
	319484	T91772		gb:yd52a10.s1 Soares fetal liver spleen	2.81	4.88
	319486	AJ382429	Hs.250799	ESTs	2.08	2.82
	319508	T99898	Hs.270104	ESTs, Moderately similar to ALU8_HUMAN A	2.80	4.39
25	319523	T69499	Hs.191184	ESTs	1.55	3.25
	319545	R83716	Hs.14355	Homo sapiens cDNA FLJ13207 fis, clone NT	1.65	1.19
	319546	R09692		gb:yd23b12.r1 Soares fetal liver spleen	5.11	8.54
	319552	AA095106	Hs.20403	ESTs	1.89	3.36
	319582	T82998	Hs.250154	hypothetical protein FLJ12973	3.48	4.82
30	319586	D78808	Hs.283683	chromosome 8 open reading frame 4	0.26	0.82
	319604	R11679	Hs.297753	vimentin	1.68	3.41
	319609	AW247514	Hs.12293	hypothetical protein FLJ21103	3.06	4.24
	319611	H14957		gb:ym19c10.r1 Soares infant brain 1NB H	2.76	4.24
	319653	AA770183	Hs.173515	uncharacterized hypothalamus protein HT0	2.51	3.55
35	319657	R19897	Hs.106604	ESTs	5.32	7.68
	319658	R13432	Hs.167481	syntrophin, gamma 1	3.35	5.00
	319661	H08035	Hs.21398	ESTs, Moderately similar to A Chain A, H	5.18	12.55
	319662	H06382	Hs.21400	ESTs	1.58	1.56
	319708	R15372	Hs.22664	ESTs	1.00	1.22
40	319742	T77668	Hs.21162	ESTs	2.48	3.13
	319748	R18178	Hs.295866	Homo sapiens mRNA; cDNA DKFZp434N1923 (f	3.02	4.85
	319772	R76633	Hs.22646	ESTs	4.36	11.61
	319788	AA321932	Hs.117414	KIAA1320 protein	2.56	3.68
45	319805	R92857	Hs.271350	likely ortholog of mouse polydom	4.63	6.56
	319812	N74880	Hs.264330	N-acylsphingosine amidohydrolase (acid c	0.60	1.32
	319834	AA071267		gb:zm61g01.r1 Stratagene fibroblast (937	0.30	0.94
	319878	T78517	Hs.13941	ESTs	3.99	6.44
	319882	AA258981	Hs.291392	ESTs	5.09	7.36
50	319912	T77559	Hs.94109	Homo sapiens cDNA FLJ13634 fis, clone PL	3.24	3.21
	319935	H79460	Hs.271722	ESTs, Weakly similar to ALU1_HUMAN ALU S	4.40	9.42
	319944	T79248	Hs.133510	ESTs	3.31	5.39
	319947	AA160967	Hs.14479	Homo sapiens cDNA FLJ14199 fis, clone NT	2.90	4.95
	319962	H06350	Hs.135056	Human DNA sequence from clone RP5-850E9	1.81	1.57
55	320007	AA336314		gb:EST40943 Endometrial tumor Homo sapie	3.42	6.29
	320018	T83263		gb:yd40h09.r1 Soares fetal liver spleen	2.77	5.14
	320030	H63789	Hs.296288	ESTs, Weakly similar to KIAA0638 protein	4.10	6.69
	320032	A1699772	Hs.292664	ESTs, Weakly similar to A46010 X-linked	3.27	3.27
	320040	AA233671	Hs.87164	hypothetical protein FLJ14001	1.81	1.64
	320047	T86564	Hs.302256	EST	3.38	7.36
60	320063	AA074108	Hs.120844	FOXJ2 forkhead factor	5.90	16.73
	320096	H58138	Hs.117915	ESTs	2.08	4.47
	320099	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	1.00	1.00
	320112	T92107	Hs.188489	ESTs	2.27	2.06
	320140	H94179	Hs.119023	SMC2 (structural maintenance of chromoso	1.00	1.00
65	320188	AW419200	Hs.172318	ESTs	1.26	1.00
	320193	AA831259	Hs.17132	ESTs	2.58	6.23
	320195	R62203	Hs.24321	Homo sapiens cDNA FLJ12028 fis, clone HE	2.85	4.53
	320199	R78659	Hs.29792	ESTs	0.40	0.94
	320203	AL049227	Hs.124776	Homo sapiens mRNA; cDNA DKFZp564N1116 (f	0.84	1.18
70	320219	AA327564	Hs.127011	tubulointerstitial nephritis antigen	1.00	1.17
	320220	AF054910	Hs.127111	teklin 2 (testicular)	0.18	1.09
	320225	AF058989	Hs.128231	G antigen, family B, 1 (prostate associa	5.26	13.75
	320231	H03139	Hs.24683	ESTs	1.59	1.93
	320260	NM_003608	Hs.131924	G protein-coupled receptor 65	1.38	4.56
75	320267	AL049337	Hs.132571	Homo sapiens mRNA; cDNA DKFZp564P016 (fr	1.09	1.92
	320268	H06019	Hs.151293	Homo sapiens cDNA FLJ10664 fis, clone NT	5.58	5.70
	320322	AF077374	Hs.139322	small proline-rich protein 3	1.41	1.01
	320325	A167978	Hs.139851	caveolin 2	-0.05	0.67
	320330	AF026004	Hs.141660	chloride channel 2	2.17	1.26
80	320339	H10807	Hs.281434	Homo sapiens cDNA FLJ14028 fis, clone HE	1.81	2.32
	320398	H16065	Hs.31286	ESTs	1.00	3.22
	320402	R22291	Hs.23368	Homo sapiens clone FLC0578 PRO2852 mRNA,	1.41	1.36
	320413	AA203711	Hs.173269	ESTs	2.31	3.61
	320432	R62786	Hs.124136	ESTs	11.25	20.78
85	320436	AA253352	Hs.293663	ESTs	2.22	3.49
	320438	W24548	Hs.5669	ESTs	3.53	8.14

	320448	AI240233	Hs.80887	v-yes-1 Yamaguchi sarcoma viral related	1.42	3.46
	320451	R26944	Hs.180777	Homo sapiens mRNA; cDNA DKFZp564M0264 (f	0.87	0.81
	320484	AA094436	Hs.296267	folistatin-like 1	0.65	1.18
5	320499	R32555	Hs.24321	Homo sapiens cDNA FLJ12028 fis, clone HE	3.44	7.15
	320514	AB007978	Hs.158278	KIAA0509 protein	6.44	13.62
	320521	N31464	Hs.24743	hypothetical protein FLJ20171	1.48	1.04
	320526	AW374205	Hs.111314	ESTs	3.66	7.87
	320527	R34672	Hs.324522	ESTs	3.16	5.63
10	320536	AA331732	Hs.137224	ESTs	2.83	5.83
	320556	AF054177	Hs.14570	hypothetical protein FLJ22530	1.28	1.00
	320564	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg	1.22	0.81
	320587	Z44524	Hs.167456	Homo sapiens mRNA full length insert cDN	1.84	2.44
	320635	R54159	Hs.80506	small nuclear ribonucleoprotein polypept	1.00	6.25
	320639	AA243258	Hs.7395	hypothetical protein FLJ23182	2.60	2.30
15	320648	N48521	Hs.26549	Homo sapiens mRNA for KIAA1708 protein,	1.00	1.53
	320651	AA489268	Hs.111334	fertilin, light polypeptide	0.14	0.79
	320664	AI904216	Hs.91251	hypothetical protein FLJ11198	5.02	8.84
	320676	AA132650	Hs.300511	ESTs	3.63	5.37
	320683	R59291	Hs.28638	ESTs, Weakly similar to unnamed protein	0.37	1.31
20	320689	AA334609	Hs.171929	ESTs, Weakly similar to A54849 collagen	1.27	1.02
	320696	AW135016	Hs.172780	ESTs	3.53	4.60
	320714	AA45591		gb:yg04a10.r1 Soares fetal liver spleen	1.06	0.85
	320727	U96044	Hs.181125	immunoglobulin lambda locus	1.35	1.49
	320771	AI793266	Hs.117176	poly(A)-binding protein, nuclear 1	0.04	0.82
25	320794	AA281993	Hs.91226	ESTs	2.96	4.33
	320822	AF100780	Hs.194679	WNT1 inducible signaling pathway protein	0.10	0.79
	320824	AF120274	Hs.194689	artemin	1.16	1.11
	320830	AJ132445	Hs.266416	claudin 14	1.06	1.75
	320843	AA317372	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (fr	1.36	1.47
30	320849	D60031	Hs.34771	ESTs	5.30	7.49
	320853	AI473796	Hs.135904	ESTs	1.00	1.00
	320896	AB002155	Hs.271580	uroplakin 1B	5.90	2.55
	320921	R94038	Hs.199538	inhibin, beta C	2.20	1.17
	320927	AI205786	Hs.213923	ESTs	0.18	1.46
35	320957	AI878933	Hs.92023	core histone macroH2A2.2	1.67	2.18
	320997	H22544		gb:yn69f11.r1 Soares adult brain N2b5HB5	3.26	3.62
	321045	W88483	Hs.293650	ESTs	2.25	4.55
	321046	H27794	Hs.269055	ESTs	2.69	4.25
	321052	AW372884	Hs.240770	nuclear cap binding protein subunit 2, 2	2.14	2.56
40	321059	AI092824	Hs.126465	ESTs	1.69	0.53
	321062	R87955	Hs.241411	Homo sapiens mRNA full length insert cDN	2.76	5.20
	321067	AF131782	Hs.241438	Homo sapiens clone 24941 mRNA sequence	4.79	7.41
	321102	AA018306		gb:za40d08.r1 Soares retina N2b4HR Homo	1.79	4.27
	321130	H43750	Hs.125494	ESTs	1.00	3.14
45	321142	AI817933	Hs.298351	ASPL protein	8.73	15.36
	321155	AA336635	Hs.99598	hypothetical protein MGC5338	3.04	5.03
	321158	AA700289		gb:yu76f11.r1 Soares fetal liver spleen	4.62	8.39
	321170	N53742	Hs.172982	ESTs	2.21	4.46
50	321199	AW385512		gb:yy56d10.s1 Soares_multiple_sclerosis_	5.69	8.01
	321206	H54178	Hs.226469	Homo sapiens cDNA FLJ12417 fis, clone MA	4.00	7.32
	321225	AL080073	Hs.251414	Homo sapiens mRNA; cDNA DKFZp564B1462 (f	4.17	4.63
	321236	AW371941	Hs.18192	Ser/Arg-related nuclear matrix protein (1.00	1.00
	321244	AF086654		gb:Homo sapiens isolate AN.1 immunoglobu	2.18	9.13
55	321270	R83560		gb:yy76c06.s1 Soares fetal liver spleen	3.80	5.26
	321317	AI937060	Hs.6298	KIAA1151 protein	1.81	1.65
	321318	AB033041	Hs.137507	KIAA1215 protein	1.00	1.00
	321325	AB033100	Hs.300646	KIAA protein (similar to mouse paladin)	0.44	0.93
	321342	AA127984	Hs.222024	transcription factor BMAL2	4.94	4.93
	321356	R93443	Hs.271770	ESTs	3.10	4.66
60	321418	AJ739161	Hs.161075	ESTs	2.28	2.54
	321420	AI368667	Hs.132743	ESTs	1.13	0.97
	321430	U05890		gb:Hsapiens (DIG3) mRNA for immunoglobu	2.42	3.35
	321453	N50080	Hs.82845	Homo sapiens cDNA: FLJ21930 fis, clone H	1.60	3.11
	321467	X13075		gb:Human 2a12 mRNA for kappa-immunoglobu	0.42	0.72
65	321468	AA514198	Hs.38540	ESTs	2.46	6.50
	321491	H70665	Hs.292549	ESTs	1.00	1.25
	321498	AW295517	Hs.255436	ESTs	3.19	6.24
	321504	W02356	Hs.268980	ESTs	2.28	3.86
	321510	AA703650	Hs.255748	ESTs	2.14	3.94
70	321513	H84972	Hs.108551	ESTs	2.78	5.37
	321516	AI382803	Hs.159235	ESTs	3.06	7.19
	321565	AI525773	Hs.266514	hypothetical protein FLJ11342	4.89	7.82
	321577	H84260		gb:ys90g04.r1 Soares retina N2b5HR Homo	1.00	1.73
	321581	AA019964	Hs.28803	ESTs	4.88	6.73
75	321582	AA143755	Hs.21858	trinucleotide repeat containing 3	1.00	2.08
	321587	H95531		gb:ys76e02.r1 Soares retina N2b4HR Homo	2.26	4.52
	321626	AA295430	Hs.96322	hypothetical protein FLJ23560	1.95	3.83
	321628	H87064	Hs.161051	ESTs, Moderately similar to ALU6_HUMAN A	0.47	1.02
	321642	AW065917	Hs.247094	ESTs	1.52	1.39
80	321669	H95404	Hs.294110	ESTs	2.17	2.45
	321687	AA625149		gb:af70c12.r1 Soares_NhiHMPu_S1 Homo sapi	4.31	6.95
	321688	H97646	Hs.123158	Homo sapiens cDNA FLJ12830 fis, clone NT	2.82	3.28
	321693	AA700017	Hs.173737	ras-related C3 botulinum toxin substrate	0.51	1.08
	321700	N55160	Hs.167260	ESTs	4.57	7.46
85	321701	AW390923	Hs.42568	ESTs	1.00	1.00

	321709	N25847	Hs.108923	RAB38, member RAS oncogene family	1.00	1.00
	321710	N35682	Hs.259743	ESTs	2.97	5.26
	321775	AI694875	Hs.202312	Homo sapiens clone N11 Ntera2D1 teratoca	1.00	1.00
	321777	AI637993	Hs.202312	Homo sapiens clone N11 Ntera2D1 teratoca	1.68	0.45
5	321779	N42729	Hs.163335	ESTs	0.90	0.90
	321829	D81993	Hs.8965	tumor endothelial marker 8	2.69	3.69
	321846	AA281594	Hs.87902	ESTs	5.11	7.64
	321879	AL109670	Hs.302809	ESTs	6.49	9.58
10	321883	AA426494	Hs.46901	KIAA1462 protein	0.28	0.95
	321899	N55158	Hs.29468	ESTs	0.39	0.95
	321911	AF026944	Hs.293797	ESTs	6.20	10.76
	321949	R49202	Hs.181694	EST	4.62	10.51
	321955	AI651866	Hs.195689	ESTs	2.89	5.47
	321956	AL110177	Hs.132882	ESTs	0.32	1.25
15	321987	AL133612	Hs.272759	KIAA1457 protein	1.00	1.83
	321991	AL133627	Hs.158923	Homo sapiens mRNA; cDNA DKFZp434K0722 (f	4.00	6.47
	322002	AA328801	Hs.84522	ESTs	2.10	3.48
	322035	AL137517	Hs.306201	hypothetical protein DKFZp564O1278	1.00	1.90
20	322044	AW340926	Hs.154679	gb:xy51b10.x1 NCI_CGAP_Lu34.1 Homo sapie	3.20	9.67
	322057	N92197	Hs.154679	synaptotagmin 1	1.55	1.07
	322060	AI341937	Hs.154679	gb:qt10e03.x1 NCI_CGAP_GC4 Homo sapiens	4.59	7.68
	322070	U80769	Hs.210322	Homo sapiens mRNA for KIAA1766 protein,	2.78	4.52
	322083	AF074982	Hs.226031	ESTs, Highly similar to KIAA0535 protein	3.10	5.52
	322091	AI819863	Hs.106243	ESTs	1.59	1.75
25	322125	R93901	Hs.117767	gb:yy16c12.r1 Soares fetal liver spleen	2.06	5.27
	322130	R98978	Hs.117767	ESTs	10.12	16.49
	322147	AF085919	Hs.114176	ESTs	0.94	0.64
	322166	AF085958	Hs.114176	gb:yr88b03.r1 Soares fetal liver spleen	4.09	6.67
30	322173	H52567	Hs.114176	gb:yl85d04.r1 Soares_pineal_gland_N3HPG	3.46	4.85
	322178	H56535	Hs.114176	gb:yl88g03.r1 Soares_pineal_gland_N3HPG	0.44	2.54
	322179	H92891	Hs.114176	gb:yl94c02.s1 Soares_pineal_gland_N3HPG	4.52	7.50
	322186	H67346	Hs.269187	ESTs	0.15	0.98
	322196	W87895	Hs.211516	ESTs	2.20	5.04
	322212	AF087995	Hs.134877	ESTs	3.42	4.84
35	322221	AI890619	Hs.179562	nucleosome assembly protein 1-like 1	0.82	2.14
	322277	AI640193	Hs.226389	ESTs	3.62	3.98
	322278	AF086283	Hs.226389	gb:zd46f01.r1 Soares_fetal_heart_NbHH19W	1.00	1.00
	322284	AI792140	Hs.49265	ESTs	0.66	2.76
40	322288	AL037273	Hs.7886	pellino (Drosophila) homolog 1	0.71	0.70
	322320	AF086419	Hs.7886	gb:zd78d03.r1 Soares_fetal_heart_NbHH19W	2.02	2.76
	322336	AA308526	Hs.76152	decorin	2.92	4.44
	322339	W17348	Hs.76152	gb:zb18c07.x5 Soares_fetal_lung_NbHL19W	8.50	11.56
	322366	AW404274	Hs.122492	hypothetical protein	0.61	1.34
45	322372	W25524	Hs.153943	ESTs	7.37	12.07
	322374	AI394663	Hs.122116	ESTs, Moderately similar to Ost2 [M.musc	4.78	10.50
	322378	AF064819	Hs.201877	DESC1 protein	1.00	1.00
	322388	AI815730	Hs.247474	hypothetical protein FLJ21032	7.09	8.49
	322416	AA223183	Hs.298442	adaptor-related protein complex 3, mu 1	3.20	5.80
	322419	AA248987	Hs.14084	ring finger protein 7	1.64	1.57
50	322425	W37943	Hs.34892	KIAA1323 protein	0.83	1.00
	322431	AA069222	Hs.141892	ESTs	3.96	5.22
	322450	AA040131	Hs.25144	ESTs	5.18	12.67
	322465	AA137152	Hs.286049	phosphoserine aminotransferase	3.41	2.23
	322467	AF116826	Hs.180340	putative protein-tyrosine kinase	1.00	1.30
55	322473	AA744286	Hs.266935	tRNA selenocysteine associated protein	1.75	2.03
	322509	T52172	Hs.302213	ESTs	1.00	2.27
	322523	W80398	Hs.193197	ESTs	2.75	5.49
	322527	AF147359	Hs.270947	gb:Homo sapiens full length insert cDNA	1.25	1.27
60	322560	AI916847	Hs.270947	ESTs	4.57	8.81
	322566	W87285	Hs.269587	ESTs	1.00	1.42
	322585	AA837622	Hs.269587	gb:zh69c01.r1 Soares_fetal_liver_spleen_	4.18	6.94
	322635	AA679084	Hs.256042	gb:zh90h08.r1 Soares_fetal_liver_spleen_	2.40	4.85
	322641	AA007352	Hs.256042	ESTs	2.94	4.64
	322653	AI828854	Hs.256538	striatin, calmodulin-binding protein	0.48	0.38
65	322664	AA011522	Hs.256538	gb:zi03g07.r1 Soares_fetal_liver_spleen_	1.92	2.18
	322687	AI110759	Hs.60843	gb:AF074566 Human fetal liver cDNA libra	4.14	6.75
	322692	AA018117	Hs.279812	potassium voltage-gated channel, shaker-	3.50	5.00
	322694	AI110872	Hs.279812	PRO327 protein	1.80	1.72
	322708	AF113674	Hs.283773	clone FLB1727	1.00	3.43
70	322712	AA021328	Hs.23507	hypothetical protein FLJ11109	3.28	3.86
	322766	AW068805	Hs.288467	Homo sapiens cDNA FLJ12280 fis, clone MA	1.63	1.53
	322770	AA045796	Hs.122682	ESTs	1.53	1.06
	322794	AI608591	Hs.38991	S100 calcium-binding protein A2	12.05	1.94
	322810	AI962276	Hs.127444	ESTs	4.09	6.90
75	322818	AW043782	Hs.293516	ESTs	1.20	1.63
	322820	AI377755	Hs.120695	ESTs	0.21	1.93
	322872	AA827228	Hs.126943	ESTs	2.04	1.63
	322882	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	5.26	1.22
	322887	AI986306	Hs.86149	phosphoinositol 3-phosphate-binding prot	2.80	2.24
80	322913	AI733737	Hs.68637	ESTs	2.38	6.61
	322926	AI825940	Hs.211192	ESTs	4.02	5.79
	322929	AI355585	Hs.146245	ESTs	0.30	1.14
	322958	AI905228	Hs.83484	SRY (sex determining region Y)-box 4	2.06	1.13
	322971	C15953	Hs.212760	hypothetical protein FLJ13649	1.18	2.00
85	322981	AA493252	Hs.159577	ESTs	2.28	2.61

	322988	C18727	Hs.171941	ESTs	0.39	2.00
	323003	AJ733859	Hs.149089	ESTs	3.28	1.00
	323013	AA134042	Hs.191451	ESTs	3.38	5.68
5	323025	AL157565	Hs.315369	Homo sapiens cDNA: FLJ23075 fis, clone L	0.06	1.10
	323032	AW244073	Hs.145946	ESTs	10.18	21.27
	323052	R21124	Hs.85573	Homo sapiens DC29 mRNA, complete cds	1.46	1.90
	323064	AL119341	Hs.49359	Homo sapiens mRNA; cDNA DKFZp547E052 (fr	3.08	5.64
	323098	AJ700025	Hs.270471	ESTs	2.31	4.49
10	323102	AL119913	Hs.163615	ESTs	5.38	11.64
	323155	AL135041		gb:DKFZp762K2310_r1 762 (synonym: hmel2)	2.38	5.56
	323176	AW071648	Hs.82101	pleckstrin homology-like domain, family	1.06	1.41
	323191	AA195600	Hs.301570	ESTs	0.73	1.24
	323225	AA205554	Hs.24790	KIAA1573 protein	5.25	11.95
15	323232	AA146722	Hs.224680	ESTs	0.45	1.35
	323266	AW003362	Hs.243886	nuclear autoantigenic sperm protein (his	1.71	1.83
	323281	AI697556	Hs.292659	ESTs	1.24	3.21
	323283	AA256014	Hs.86682	Homo sapiens cDNA: FLJ21578 fis, clone C	12.68	15.05
	323314	AA226310	Hs.191501	ESTs	4.42	9.61
20	323316	AL134620	Hs.280175	ESTs	2.98	5.93
	323334	AL36501	Hs.77273	ras homolog gene family, member A	1.98	3.30
	323338	R74219	Hs.23348	S-phase kinase-associated protein 2 (p45	1.62	1.00
	323348	AA233056	Hs.191518	ESTs	1.00	1.07
	323351	AA704103	Hs.24049	ESTs	1.43	1.68
25	323359	AA234172	Hs.137418	ESTs	0.34	1.18
	323360	AA716061	Hs.161719	ESTs	3.01	3.71
	323405	AW139550	Hs.115173	ESTs	1.90	8.81
	323420	AJ672386	Hs.263780	ESTs	0.29	1.01
	323434	AW081455	Hs.120219	ESTs	2.27	1.92
30	323445	AA253103	Hs.135569	ESTs, Weakly similar to NEUROD [H.sapien	0.43	0.80
	323449	AA282865	Hs.284153	Fanconi anemia, complementation group A	3.19	3.85
	323492	H00978	Hs.20887	hypothetical protein FLJ10392	2.70	3.20
	323501	AA182461	Hs.84520	ESTs	2.04	3.31
	323505	AI652287		gb:EST382593 MAGE resequences, MAGK Homo2.21		3.08
35	323515	AA282274	Hs.256083	ESTs	2.69	3.40
	323541	AI185116	Hs.104613	RP42 homolog	1.20	1.09
	323545	AI814405	Hs.224569	ESTs	1.25	1.55
	323635	R63117	Hs.9691	Homo sapiens cDNA: FLJ23249 fis, clone C	0.27	0.72
40	323675	AA984759	Hs.272168	tumor differentially expressed 1	3.70	5.80
	323678	AL042121	Hs.20880	ESTs	3.33	5.10
	323691	AA317561	Hs.145599	ESTs	1.00	1.00
	323693	AW297758	Hs.249721	ESTs	2.01	1.54
	323746	AW298611	Hs.12808	MARK	4.11	5.53
45	323774	AA329806	Hs.321056	Homo sapiens mRNA; cDNA DKFZp586F1322 (f	2.06	3.70
	323856	AA355264	Hs.267604	hypothetical protein FLJ10450	3.42	8.13
	323857	T18988	Hs.293668	ESTs	5.97	12.51
	323870	AA341774	Hs.129212	ESTs	3.17	4.52
	323876	AL042492	Hs.147313	ESTs	0.36	1.00
50	323885	AA344308	Hs.128427	Homo sapiens BAC clone RP11-335J18 from	2.31	3.33
	323911	AL043212	Hs.92550	ESTs	4.38	5.41
	323919	AA862973	Hs.220704	ESTs	5.80	10.20
	323972	AI869964	Hs.182906	ESTs	3.10	5.14
	324005	AA610011	Hs.208021	ESTs	5.34	10.07
55	324036	AI472078	Hs.303662	ESTs	1.00	5.03
	324055	AA526794	Hs.128644	ESTs	0.86	1.00
	324063	AW292740	Hs.272813	dual oxidase 1	0.45	0.91
	324072	AA381829		gb:EST94855 Activated T-cells I Homo sap	2.82	5.12
	324092	AW269931	Hs.202473	Homo sapiens cDNA: FLJ22278 fis, clone H	2.40	2.52
	324095	AW377983	Hs.298140	Homo sapiens cDNA: FLJ22502 fis, clone H	1.32	4.30
60	324129	AI381918	Hs.285833	Homo sapiens cDNA: FLJ22135 fis, clone H	1.40	1.77
	324132	AW504860	Hs.288836	hypothetical protein FLJ12673	4.24	6.21
	324214	AA412395	Hs.225740	ESTs	6.96	10.69
	324227	AA295552	Hs.28631	Homo sapiens cDNA: FLJ22141 fis, clone H	0.81	0.53
	324266	AL047634	Hs.231913	ESTs	2.42	4.05
65	324275	AA429088	Hs.98523	ESTs	3.62	5.38
	324281	AL048026	Hs.124675	ESTs, Weakly similar to T14742 hypotheti	0.14	0.70
	324290	AA432032	Hs.304420	ESTs	3.71	4.34
	324303	AL118754		gb:DKFZp761P1910_r1 761 (synonym: hamy2)	0.95	0.91
70	324312	AI198841	Hs.128173	ESTs	4.06	5.91
	324325	AL138153	Hs.300410	ESTs	5.88	8.25
	324338	AL138357	Hs.145078	regulator of differentiation (in S. pombe	0.87	1.25
	324341	AW197734	Hs.99807	ESTs, Weakly similar to unnamed protein	1.28	1.00
	324343	AW452016	Hs.293232	ESTs	2.54	3.46
	324371	AA452305	Hs.270319	ESTs	5.85	8.36
75	324382	AW502749	Hs.24724	MFH-amplified sequences with leucine-ric	0.76	1.64
	324384	AA453396	Hs.127656	KIAA1349 protein	2.88	5.69
	324385	F28212	Hs.284247	KIAA1491 protein	1.81	1.99
	324388	AI924963	Hs.306206	hypothetical protein FLJ11215	1.00	1.00
	324432	AA464510	Hs.152812	ESTs	2.73	2.17
	324497	AW152624	Hs.136340	ESTs, Weakly similar to unnamed protein	0.71	1.90
80	324510	AI148353	Hs.287425	Homo sapiens cDNA FLJ11569 fis, clone HE	1.00	1.00
	324580	AA492588		gb:ng99c08.s1 NCL CGAP_Thy1 Homo sapiens	2.18	3.50
	324582	AA506935	Hs.132036	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.96	11.36
	324633	AA572994	Hs.325489	ESTs	2.92	4.22
	324640	AW295832	Hs.134798	ESTs, Moderately similar to TTL MOUSE TU	5.48	11.74
85	324675	AW014734	Hs.157969	ESTs	0.39	0.73

	324699	AW504732	Hs.21275	hypothetical protein FLJ11011	0.93	0.93
	324747	AA603532	Hs.130807	ESTs	1.57	1.81
	324748	AA657457	Hs.292385	ESTs	1.55	1.34
	324801	AI819924	Hs.14553	sterol O-acyltransferase (acyl-Coenzyme	1.00	6.56
5	324804	AI692552		gb:wd73f12x1 NCI_CGAP_Lu24 Homo sapiens	1.00	7.53
	324828	AA843926	Hs.124434	ESTs	2.00	3.25
	324855	AW152305	Hs.122364	ESTs	2.74	3.43
	324866	AI541214	Hs.46320	Small proline-rich protein SPRK (human,	1.07	0.95
	324871	AW297755	Hs.271923	Homo sapiens cDNA: FLJ22785 fis, clone K	1.68	1.21
10	324886	AA806794	Hs.131511	ESTs	2.56	5.61
	324889	D31010		gb:HUML12147 Human fetal lung Homo sapie	2.20	4.65
	324948	AW383618	Hs.265459	ESTs, Moderately similar to ALU2_HUMAN A	5.28	7.05
	324953	AI264628	Hs.125428	ESTs	3.37	5.51
	324958	AA625076	Hs.132892	protocadherin 20	5.12	9.81
15	324988	T06997	Hs.121028	hypothetical protein FLJ10549	2.52	1.08
	325024	F13254	Hs.78672	laminin, alpha 4	5.24	10.22
	325105	H97109	Hs.105421	ESTs	1.00	1.00
	325108	AA401863	Hs.22380	ESTs	1.99	2.14
	325114	D83901	Hs.315562	ESTs	2.73	3.17
20	325146	AI064690	Hs.171176	ESTs	1.86	3.41
	325149	D61117	Hs.187646	ESTs	0.42	0.93
	325187	AI653682	Hs.197812	ESTs	6.50	11.31
	325228				6.18	15.75
	325235				2.64	4.12
25	325328				2.87	4.42
	325340				0.29	0.33
	325367				16.56	24.29
	325373				0.63	1.22
	325389				0.88	1.05
30	325436				5.75	14.14
	325471				8.46	17.82
	325498				3.32	6.42
	325557				5.51	8.28
	325559				7.48	21.40
35	325560				4.08	6.25
	325569				4.20	5.24
	325585				1.10	1.13
	325587				1.00	1.00
	325597				2.98	13.40
40	325639				0.78	0.78
	325639				0.46	0.66
	325685				0.95	1.55
	325686				4.48	9.20
	325735				0.59	0.88
45	325739				2.42	6.61
	325740				7.88	9.83
	325792				4.74	7.18
	325819				2.02	2.64
	325883				7.78	15.98
50	325895				2.04	10.60
	325925				4.18	7.36
	325932				3.66	9.03
	325941				0.61	0.80
	325969				4.88	7.42
55	325971				0.55	1.07
	326025				7.21	14.72
	326046				3.60	5.99
	326099				1.27	1.06
	326108				3.27	5.70
60	326163				0.45	1.11
	326165				0.13	0.45
	326189				5.60	9.00
	326204				7.00	12.01
	326230				1.00	8.09
65	326274				9.86	15.35
	326360				0.52	0.77
	326393				1.00	1.42
	326505				1.24	5.84
	326515				9.20	13.49
70	326589				2.77	4.01
	326592				2.01	2.53
	326605				1.00	1.00
	326692				1.00	1.31
	326693				0.19	0.65
75	326720				2.34	7.20
	326742				0.25	0.83
	326770				3.09	4.56
	326818				2.08	3.45
	326936				0.41	1.70
80	326964				2.02	3.80
	326983				1.09	1.20
	326991				1.00	8.04
	327036				3.05	4.22
	327040				3.55	6.31
85	327053				1.59	1.40
	327075					

WO 02/086443

PCT/US02/12476

	327085	2.50	12.57
	327130	5.38	8.04
	327156	3.74	6.58
5	327220	1.28	1.54
	327224	6.56	12.91
	327268	2.61	5.40
	327321	2.42	3.11
	327332	6.62	10.58
10	327361	2.69	4.41
	327377	2.04	6.72
	327396	2.61	4.50
	327414	1.00	8.01
	327442	5.91	9.65
15	327467	6.58	18.01
	327473	3.79	7.48
	327483	4.08	8.87
	327562	0.68	2.86
	327568	1.00	2.00
20	327606	2.06	3.61
	327611	5.90	14.26
	327642	4.06	8.74
	327654	1.05	2.08
	327734	1.00	1.00
25	327775	1.46	11.79
	327796	3.47	5.65
	327840	3.26	6.64
	327940	5.84	15.58
	327984	0.36	1.50
30	328004	1.87	1.42
	328021	0.42	0.59
	328068	2.83	4.68
	328100	3.04	5.39
	328101	3.54	5.20
35	328113	0.72	0.91
	328157	5.58	5.16
	328196	5.76	11.13
	328197	5.98	10.58
	328264	3.11	4.88
40	328299	2.20	3.06
	328342	1.49	1.94
	328365	1.00	1.00
	328369	4.40	7.36
	328381	1.86	4.93
45	328451	5.51	7.56
	328481	0.13	0.72
	328500	2.71	3.97
	328530	5.41	7.62
	328600	3.14	10.68
50	328608	4.56	8.17
	328616	2.24	11.91
	328623	3.04	5.46
	328632	0.70	1.19
	328664	3.48	6.80
55	328666	10.42	26.47
	328698	9.68	14.56
	328700	2.74	10.22
	328708	0.15	0.57
	328735	6.23	8.91
60	328743	3.62	6.54
	328806	0.22	0.78
	328861	3.68	10.54
	328908	5.42	16.36
	328933	2.02	5.29
65	328934	1.73	4.45
	328949	3.34	5.41
	329005	2.88	7.26
	329011	2.52	3.72
	329033	1.00	1.03
70	329037	5.07	8.16
	329067	1.98	2.41
	329134	2.24	3.25
	329157	2.30	11.04
	329178	2.64	5.02
75	329192	6.41	15.27
	329194	0.31	0.79
	329204	1.60	3.75
	329224	2.99	6.11
	329228	0.83	0.83
80	329288	0.63	1.01
	329337	1.00	1.00
	329541	0.76	1.68
	329560	1.34	2.02
	329588	1.68	2.22
85	329643	4.18	11.77
	329703	1.00	1.00

	329764			5.78	15.50
	329816			2.09	5.44
	329860			3.13	10.77
	329993			7.83	14.21
5	330020			5.58	13.12
	330036			3.32	5.57
	330052			4.31	7.97
	330085			1.34	1.76
	330088			4.70	12.46
10	330093			0.44	1.06
	330100			3.47	4.83
	330105			2.14	3.61
	330107			3.17	6.87
	330120			5.61	11.89
15	330123			4.50	12.74
	330208			1.55	7.62
	330263			13.10	23.38
	330300			2.81	4.98
	330313			3.00	4.41
20	330366			0.67	0.76
	330372			4.76	11.82
	330385	AA449749	Hs.182971	2.14	2.15
	330397	D14659	Hs.154387	0.40	1.15
	330468	L10343	Hs.112341	1.11	0.94
25	330472	L24203	Hs.82237	1.67	1.17
	330478	L38486	Hs.296049	0.46	1.07
	330493	M27826	Hs.267319	1.07	0.95
	330495	M31328	Hs.71642	0.97	0.96
	330506	M61906	Hs.6241	0.17	3.66
30	330512	M80563	Hs.81256	0.60	1.06
	330537	U19765	Hs.2110	2.81	2.07
	330547	U32989	Hs.183671	3.91	1.49
	330551	U39840	Hs.299867	1.15	1.03
	330568	U56244	(NONE)	2.83	4.79
35	330599	U90437	gb:Human RP1 homolog mRNA, 3'UTR region	2.08	1.54
	330601	U90916	Hs.82845	0.89	1.35
	330605	X02419	Hs.77274	1.87	1.55
	330609	X04741	Hs.76118	1.83	1.30
	330617	X53587	Hs.85266	1.54	1.15
40	330630	X78669	Hs.79088	1.39	1.19
	330644	Y07755	Hs.38991	3.83	1.13
	330650	Z68228	Hs.2340	1.25	0.95
	330660	AA347868	Hs.139293	15.50	29.07
	330692	AA017045	Hs.6702	1.00	1.00
45	330707	AA133891	Hs.293690	0.20	1.35
	330715	AA233707	Hs.11571	0.12	1.40
	330717	AA233926	Hs.52620	6.62	5.42
	330722	AA243560	Hs.34382	1.40	1.65
	330740	AA297746	Hs.22654	0.27	2.04
50	330742	AA400979	Hs.25691	0.44	0.90
	330744	AA406142	Hs.12393	0.71	3.23
	330751	AA428286	Hs.29643	1.66	1.52
	330760	AA448663	Hs.30469	0.52	0.90
	330763	AA450200	Hs.274337	0.37	0.97
55	330786	D60374	Hs.49136	0.78	0.84
	330790	T48536	Hs.105807	0.23	3.17
	330814	AA015730	Hs.265398	0.37	2.07
	330827	AA040332	Hs.12744	1.60	1.00
	330844	AA053037	Hs.66803	0.93	1.16
60	330901	AA157818	Hs.267319	1.02	1.03
	330931	F01443	Hs.284256	0.24	0.88
	330952	H02855	Hs.29567	0.08	1.31
	330961	H10998	Hs.7164	1.29	1.26
	330968	H16568	Hs.23748	0.48	0.96
65	331014	H98597	Hs.30340	0.29	0.74
	331046	N66563	Hs.191358	0.99	8.56
	331060	N75081	Hs.157148	1.24	1.00
	331099	R36671	Hs.83937	0.75	1.03
	331108	R41408	Hs.21583	1.00	2.75
70	331131	R54797	gb:yg87b07.s1 Soares infant brain 1N1B H	6.04	10.68
	331135	R61398	Hs.4197	0.80	0.96
	331170	T23451	Hs.159293	2.63	4.29
	331180	T32446	Hs.6640	1.78	2.71
	331183	T40769	Hs.8469	1.00	3.01
75	331203	T82310	(NONE)	1.70	3.80
	331271	AA059347	Hs.82226	1.20	3.19
	331306	AA252079	Hs.63931	0.31	1.30
	331327	AA281076	Hs.109221	2.09	2.41
	331341	AA303125	Hs.23240	0.72	2.43
80	331359	AA416979	Hs.46901	0.09	0.91
	331363	AA421562	Hs.91011	1.02	0.87
	331378	AA448881	Hs.49282	1.03	1.23
	331384	AA456001	Hs.93847	1.40	1.00
	331402	AA505135	Hs.44037	1.80	3.93
85	331422	F10802	Hs.163628	1.65	1.89

	331490	N32912	Hs.26813	CDA14	2.48	1.73
	331531	N51343		gbxyz15g04.s1 Soares_multiple_sclerosis	0.98	1.68
	331547	N54811		gbxod74f04.s1 NCI_CGAP_Ov2 Homo sapiens	3.80	5.75
5	331578	N67960	Hs.249989	ESTs	0.11	0.67
	331589	N71027	Hs.152618	ESTs	1.09	1.38
	331608	N89861	Hs.112110	PTD007 protein	0.93	0.76
	331614	N92293	Hs.240272	EST	0.17	1.34
	331668	W69707	Hs.58030	EST	2.24	3.82
10	331671	W72033	Hs.194695	ras homolog gene family, member 1	1.00	1.24
	331676	W79834	Hs.58559	ESTs, Weakly similar to rholekin (Mmusc	0.08	1.07
	331681	W85712	Hs.119571	collagen, type III, alpha 1 (Ehlers-Danl	8.72	4.27
	331692	W93592	Hs.152213	wingless-type MMTV integration site fami	0.94	0.54
	331717	AA190888	Hs.153881	Homo sapiens NY-REN-62 antigen mRNA, par	1.57	1.34
15	331718	AA191404	Hs.104072	ESTs	6.80	11.77
	331811	AA404500	Hs.301570	ESTs	1.10	1.00
	331820	AA405970	Hs.97996	transcription termination factor, mitoc	0.73	0.59
	331831	AA412031	Hs.97901	EST	2.77	4.08
	331852	AA418988	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	0.23	0.93
20	331943	AA453418	Hs.21275	hypothetical protein FLJ11011	0.36	1.88
	331969	AA460702	Hs.82772	collagen, type XI, alpha 1	1.00	1.00
	331990	AA478102	Hs.139631	ESTs	3.04	3.87
	332002	AA482009	Hs.105104	ESTs	1.19	0.78
	332027	AA489671	Hs.65641	hypothetical protein FLJ20073	1.27	1.03
25	332029	AA489697	Hs.145053	ESTs	0.30	1.62
	332033	AA489840	Hs.251014	EST	2.30	3.70
	332048	AA496019	Hs.201591	ESTs	0.17	0.52
	332071	AA598594	Hs.205293	KIAA1211 protein	1.35	1.23
30	332074	AA599012		gbxae41e11.s1 Gessler Wilms tumor Homo s	0.19	2.00
	332083	AA600200	Hs.155546	KIAA1080 protein; Golgi-associated, gamm	0.31	1.18
	332085	AA600353	Hs.173933	nuclear factor I/A	0.30	1.50
	332125	AA609861	Hs.312447	ESTs	0.22	0.62
	332177	F10812	Hs.101433	ESTs	8.21	18.03
	332180	H03348	Hs.7327	claudin 1	2.27	1.57
35	332185	H10356	Hs.101689	ESTs	0.09	1.18
	332203	H49388	Hs.317769	EST	8.05	5.02
	332232	N48891	Hs.101915	Stargardt disease 3 (autosomal dominant)	0.78	0.85
	332240	N54803	Hs.324267	ESTs, Weakly similar to putative p150 [0.96	1.23
	332261	N70294	Hs.269137	ESTs	2.40	3.74
40	332275	R08838	Hs.26530	serum deprivation response (phosphatidyl	0.27	0.75
	332280	R38100	Hs.146381	RNA binding motif protein, X chromosome	0.39	1.88
	332299	R69250	Hs.21201	nectin 3; DKFZP566B0846 protein	5.24	12.76
	332304	R74041	Hs.101539	ESTs	1.44	3.18
	332314	T25862	Hs.101774	hypothetical protein FLJ23045	0.68	1.32
45	332384	M11433	Hs.101850	retinol-binding protein 1, cellular	1.71	0.88
	332434	N75542	Hs.289068	Homo sapiens cDNA FLJ11918 fis, clone HE	0.43	0.86
	332445	T63781	Hs.11112	ESTs	0.68	1.00
	332453	L00205	Hs.111758	keratin 6A	31.54	1.00
	332458	M33493	Hs.250700	trypsin beta 1	0.51	1.00
50	332504	AA053917	Hs.15106	chromosome 14 open reading frame 1	0.79	1.24
	332525	M17252	Hs.278430	cytochrome P450, subfamily XXIA (steroid	0.98	1.70
	332530	M31682	Hs.1735	inhibin, beta B (activin AB beta polypep	0.88	0.66
	332535	N20284	Hs.19280	cysteine-rich motor neuron 1	0.22	1.46
	332539	AA412528	Hs.20183	ESTs, Weakly similar to AF164793 1 prote	0.93	1.49
55	332559	M13955	Hs.166189	cytokeratin 2	0.35	1.13
	332563	N92824	Hs.274407	protease, serine, 16 (thymus)	1.00	1.00
	332565	AA234896	Hs.25272	E1A binding protein p300	0.36	1.05
	332594	AA279313	Hs.3239	methyl CpG binding protein 2 (Rett syndr	0.53	0.59
	332634	S38953	Hs.283750	tenascin XA	0.38	1.16
60	332638	AA283034	Hs.50640	JAK binding protein	1.00	1.70
	332640	AA417152	Hs.5101	protein regulator of cytokinesis 1	6.15	1.16
	332654	AA001296	Hs.288217	hypothetical protein MGC2941	1.50	2.73
	332665	AA223335	Hs.63788	propionyl Coenzyme A carboxylase, beta p	1.20	0.91
	332692	AA496035	Hs.247926	gap junction protein, alpha 5, 40kD (con	0.17	1.12
65	332716	L00058	Hs.79070	v-myc avian myelocytomatosis viral oncog	1.00	1.44
	332736	L13773	Hs.114765	myeloid/lymphoid or mixed-lineage leukem	1.00	1.81
	332758	X93921	Hs.296938	dual specificity phosphatase 7	0.53	0.78
	332781	AA233258	Hs.247112	hypothetical protein FLJ10902	1.44	1.56
	332792				1.70	1.19
70	332816				1.85	2.47
	332858				1.04	1.57
	332906				3.48	8.04
	332911				1.00	1.00
	332912				1.06	4.40
75	332922				1.00	1.00
	332956				0.42	0.88
	332959				1.96	6.34
	332982				0.56	0.99
	332984				0.30	0.78
80	332998				1.47	2.01
	333058				0.47	1.38
	333097				2.14	3.19
	333121				2.76	3.70
	333122				1.92	1.21
85	333123				1.85	1.39
	333138				0.47	0.52

	333139	1.88	0.84
	333140	0.21	0.64
	333221	1.51	1.11
5	333260	0.75	1.01
	333380	6.68	15.75
	333387	4.56	12.61
	333512	5.05	8.01
	333524	2.28	3.98
10	333585	2.31	1.53
	333603	2.23	1.17
	333604	2.51	1.58
	333618	0.52	0.98
	333627	1.44	1.36
15	333628	1.90	1.90
	333650	1.85	2.10
	333678	1.85	2.35
	333750	2.18	5.67
	333763	1.99	2.60
20	333767	1.02	0.96
	333768	1.78	1.65
	333769	2.15	2.13
	333772	1.46	2.53
	333777	1.00	1.42
25	333846	2.99	4.50
	333884	0.47	0.94
	333887	0.50	1.00
	333891	0.43	0.89
30	333892	0.51	0.91
	333904	0.26	1.13
	333906	0.55	0.98
	333948	1.70	2.15
	333954	0.37	1.09
	333966	8.10	14.30
35	333968	0.63	1.38
	334061	4.24	12.30
	334094	1.30	12.03
	334113	4.55	8.63
	334161	0.82	1.59
40	334183	0.47	0.76
	334187	1.36	3.70
	334219	0.69	1.04
	334222	1.88	1.70
	334223	4.72	3.14
45	334239	0.79	0.62
	334255	0.45	1.10
	334333	1.00	3.56
	334378	3.98	5.76
	334382	1.50	1.31
50	334492	3.59	4.75
	334562	5.94	15.40
	334588	8.14	19.53
	334616	1.55	1.56
	334633	5.16	8.07
55	334648	0.59	2.13
	334787	3.70	7.15
	334866	8.13	10.60
	334891	0.32	1.14
	334933	1.00	3.84
60	334934	4.01	7.43
	334945	1.04	2.96
	334967	0.29	1.14
	334990	1.50	1.39
	335015	5.88	18.65
65	335093	0.55	1.75
	335120	4.31	8.01
	335125	0.38	1.97
	335179	1.24	1.98
	335188	0.46	1.47
70	335211	1.61	1.42
	335288	0.73	0.97
	335289	0.20	0.26
	335361	2.18	1.58
	335379	0.50	0.71
75	335414	3.64	14.94
	335416	2.93	3.98
	335496	0.96	0.91
	335497	1.71	1.92
	335548	1.15	2.40
80	335551	3.22	10.54
	335553	3.42	4.89
	335586	5.50	12.75
	335619	2.99	3.07
	335620	3.80	8.29
85	335621	0.28	0.57
	335682	0.46	1.17

WO 02/086443

PCT/US02/12476

	335686	2.55	3.81
	335755	2.24	1.07
	335784	0.20	0.97
5	335814	1.13	1.48
	335815	2.45	3.51
	335823	1.00	4.16
	335835	0.49	1.70
	335851	1.66	1.39
10	335868	2.98	6.43
	335896	0.98	0.99
	335936	12.10	21.93
	335948	1.00	1.64
	335983	1.00	4.21
	335995	0.37	1.17
15	336021	1.04	0.84
	336034	11.40	23.54
	336038	1.19	1.21
	336066	0.54	1.63
20	336107	0.95	0.70
	336205	3.13	6.29
	336275	3.20	10.10
	336292	2.34	3.09
	336331	1.00	1.00
	336419	0.65	0.79
25	336632	2.33	2.16
	336633	2.55	2.23
	336634	2.19	2.03
	336635	2.69	2.48
30	336636	2.13	1.83
	336637	2.43	2.24
	336638	2.31	2.03
	336659	0.60	1.31
	336675	0.31	1.18
35	336684	1.50	1.14
	336694	4.74	7.10
	336716	4.43	6.37
	336721	2.20	0.74
	336798	1.64	2.14
40	336900	6.14	12.73
	336948	1.00	1.00
	337028	1.30	2.09
	337043	4.01	11.53
	337046	1.67	1.84
45	337054	2.78	7.35
	337128	7.20	16.14
	337162	3.45	5.34
	337183	5.72	11.41
	337184	3.72	5.90
50	337192	1.27	1.06
	337194	1.88	1.68
	337229	0.22	1.03
	337268	1.00	3.31
	337299	3.23	5.14
55	337325	2.76	3.72
	337389	5.80	10.42
	337493	2.06	6.30
	337497	7.88	20.29
	337500	3.80	4.48
60	337549	1.66	2.31
	337603	1.27	8.54
	337605	5.76	7.16
	337671	0.73	0.97
	337755	1.54	0.92
65	337786	5.07	9.73
	337809	6.18	12.87
	337862	3.78	12.97
	337871	2.66	8.16
	337958	0.26	1.34
70	338008	1.48	1.12
	338033	2.38	14.59
	338083	0.65	2.16
	338110	1.00	1.61
	338112	5.86	8.25
75	338145	1.70	1.97
	338148	8.07	18.19
	338158	1.30	4.55
	338161	2.58	3.57
	338179	1.00	1.00
80	338182	3.32	4.63
	338189	1.00	3.34
	338197	0.99	1.69
	338199	4.58	7.62
	338215	6.01	15.85
85	338279	0.53	0.95
	338316	20.58	38.66

	338322	3.23	7.39
	338357	4.10	11.39
	338359	10.12	21.59
5	338366	0.69	1.02
	338374	0.40	1.18
	338414	0.47	1.06
	338418	6.12	13.85
	338469	3.09	5.11
10	338501	6.28	10.32
	338506	6.97	12.41
	338523	3.10	5.84
	338549	1.70	2.70
	338561	0.79	0.81
	338662	1.72	1.46
15	338671	0.17	0.91
	338676	2.10	15.86
	338726	1.20	1.09
	338779	0.12	0.57
20	338804	0.99	1.67
	338836	1.00	1.00
	338871	4.30	9.81
	338872	5.02	12.81
	338879	0.23	1.12
25	338937	6.55	12.26
	338966	1.76	5.42
	338993	1.00	2.40
	339047	5.26	10.81
	339100	5.10	6.88
30	339114	1.00	1.70
	339121	1.00	3.75
	339170	10.36	19.67
	339229	4.08	13.48
	339264	2.64	3.83
35	339293	1.73	1.94

TABLE 8B shows the accession numbers for those Pkeys in Table 8A lacking unigenelD's. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

	Pkey:	Unique Eos probeset identifier number	
	CAT number:	Gene cluster number	
	Accession:	Genbank accession numbers	
45	Pkey	CAT number	Accessions
	322044	187363_1	AW340926 AA249063 N86075
50	322060	44320_1	AI341937 AW003063 U34725 AA904742
	321430	42705_1	X57414 X57415
	321467	43034_1	X13075 X13076
	322125	46779_1	R93901 AF075073 R93902
	322166	46861_1	H69434 AF085958 H69846
55	322173	46873_1	H52567 H52557 AF085970 H52164
	322178	46882_1	H56535 AF085980 H56712
	322179	46885_1	H92891 AF085982 H92777
	321577	1615102_1	H84849 H84252 H84260 H86664 H85320
	321587	1615333_1	H95531 H95521 H84529
60	313723	111953_1	AA070412 AA102346 AA081885
	320997	627492_1	H22544 H46842 AI204929
	322278	47271_1	W69304 AF086283 W69200
	321687	218439_1	AA625149 AA313030 AA313052 H97463
	313883	129439_1	AA665069 AA135130 AA484059 AA102419 AW877765
65	322320	47422_1	W79150 AF086419
	322339	814584_1	AI688646 AJ734214 W17348
	314648	293660_1	AW979268 AA878419 AA431342 AA431628
	300201	682222_1	AI308300 AI306296
	306897	25195_2	AI093967
70	323155	979809_1	AL120701 AL135041 AL121524
	322527	38927_1	AF147359 T58511 T58560
	322585	473768_2	W88919 W89125
	300362	1574395_1	Z42308 H23514
	322635	82296_1	AA005129 AA679084 AA694399
75	322664	85042_1	AA011522 AA702641 AA011691 AA330797
	315454	380580_1	AI239464 AI239473 AA625812 AI208703
	322687	37372_1	AF074666 AI110759 AF090902
	314852	327472_1	AI903735 AA491283 AI694953 AW976903 AA761362
	307783	697809_1	AI347274 AW844024
80	324072	269032_1	AA381722 AA381829 AW963906 AW963902 AA381242
	300627	221345_1	AA488472 W27363 AA317053 BE082689 AW967036 BE079872
	323505	196389_1	AW970512 AA280251 AI652287 BE466438 AI650725 AA551854 AA281574 AW571481
	315791	403558_1	AA677034
	324303	233842_1	AL118754 AA333202 H38001
85	316519	442885_1	AA847835 AA768376
	300926	333127_1	AA504860 AA504911

	324580	328264_1	AA492588 AA492498 AA492571
	301882	275087_1	T78054 T79888 AA398185
	324804	398093_1	AJ692552 AJ393343 AJ800510 AJ377711 F24263 AA661876
5	324889	1515978_1	D31010 D30991 D31168 D31166 D31465
	302697	43219_1	AJ001409 AJ001410
	302711	45419_1	L08442 D51348
	302742	458_39	L12051
	318499	364430_1	T25451 AA585296 AA585305
10	310624	34624_4	U88896 U88898 AA916056 T03285 AJ341594 AJ359534 AJ634031 U88897
	302847	458_105	X98941 X98942 X98943 X98953 X98949
	304122	77271_5	H28966
	303598	270283_1	AA382814 AA402411 AA412355
	311409	837264_1	AJ698839 AJ909260 AJ909259
	312094	797889_1	Z78390 T97427
15	319312	1540116_1	Z45481 F12393 T74437
	319407	1688823_1	R05329 R01555 R08276
	319425	1689571_1	T82930 R02424 T85145
	320007	229683_1	AA336314 T82938 AA327744 AW967388 AA639967 T10753
20	320018	1815987_1	T83263 T85731 T85730
	319484	1691553_1	T91772 R07257 R07098
	318865	1535937_1	H10818 F07831 Z43072
	312220	1671607_1	N74613 T98756 T98589
	319546	243305_1	R09692 R09414 AA346353
	312389	902067_1	AJ863140 W80703 R43474
25	319611	1566863_1	H14957 R56522 R11908
	312437	291472_1	BE080180 AW827313 AW231970 AA995028 AA428584 AW872716 AW892508 AW854593 AA578441 AW975234 AA664937 AA984131
			AA528743 AA552874 AA564758 AW053245 AJ267534 AW070190 AW893483 AA770330 AA906928 AA906582 AA758746 AA551717
			AW063311 AA429538
30	311896	579182_1	AW206447 AJ248530 AJ084433 AJ400976 R16553
	319834	112523_1	AA071267 T65940 T64515 AA071334
	321102	80531_1	AA018306 H38925 AA001221
	321158	410938_1	H79670 H47798 AA700289
	321199	212379_1	N34524 AA305071 AW954803 AA502335 AJ433430 AJ203597 AW026670 AW265323 AW850787 AA317554 AW993643 AW835572
35			AW385512 AJ334966 W32951 H62656 H53902 R88904 AW835732
	305528	28832_3	AA769156
	321270	1662057_1	N59537 N78278 R83560
	314126	177666_1	AA226431 AA226569 AA488748
	320714	743644_1	R91883 AJ445591
40	306442	AA976899	
	306446	AA977348	
	306458	AA978186	
	306510	AA988546	
	306557	AA994530	
45	306572	AA995686	
	306582	AA996248	
	306656	AJ004024	
	306686	AJ015615	
	306751	AJ032589	
50	308011	AJ439473	
	306892	AJ092465	
	308106	AJ476803	
	308154	AJ500600	
	306956	AJ125111	
55	306958	AJ125152	
	308213	AJ557041	
	308216	AJ557135	
	308219	AJ557246	
	308588	AJ718299	
60	308599	AJ719893	
	308643	AJ745040	
	308673	AJ760864	
	308697	AJ767143	
	308778	AJ811109	
65	308808	AJ818289	
	308875	AJ832332	
	308886	AJ833240	
	308898	AJ858845	
	308966	AJ870704	
	308979	AJ873111	
70	303011	41689_1	AF090405 AF090407 AF090406
	303077	44060_1	AF163305 AF163307 AF163303
	305016	AA626876	
	305034	AA630128	
75	305072	AA641012	
	305148	AA654070	
	305190	AA665955	
	303978	AW513315	
	303990	AW515465	
	303998	AW516449	
80	303999	AW516611	
	305235	AA670480	
	305312	AA700201	
	305413	AA724659	
	305447	AA737856	
85	321244	29327_1	AF068654 AF068656 AF068655

WO 02/086443

PCT/US02/12476

	305614	AA782866	
	305637	AA806124	
	305639	AA808138	
5	305650	AA807709	
	305690	AA813477	
	305728	AA828209	
	305759	AA835353	
	305792	AA845256	
10	307041	AI144243	
	307091	AI167439	
	307181	AI189251	
	305901	AA872968	
	305910	AA875981	
15	307415	AI242118	
	307426	AI243364	
	307517	AI275055	
	307551	AI281556	
	307561	AI282207	
20	307608	AI290295	
	307691	AI318285	
	307730	AI336092	
	307760	AI342387	
	307764	AI342731	
	307796	AI350556	
25	309045	AI910902	
	309051	AI911975	
	307807	AI351799	
	307808	AI351826	
	307820	AI355761	
30	307852	AI365541	
	309122	AI928178	
	309164	AI937761	
	309177	AI951118	
35	307902	AI380462	
	309299	AW003478	
	309303	AW004823	
	309476	AW129368	
	309532	AW151119	
40	309747	AW264889	
	309769	AW272346	
	309799	AW276964	
	309866	AW299916	
	302679	311853_1	H65022 AA186889
45	309923	AW340684	
	309928	AW341418	
	309931	AW341683	
	309933	AW341936	
	302705	31765_1	U09060 U09061
50	302789	34161_1	AJ245057 AJ245070
	304006	AW517947	
	304024	T03036	
	304026	T03160	
	304028	T03266	
55	304046	T54803	
	304061	T61521	
	304063	T62536	
	302802	34487_1	Y08250 Y08245
	304114	R78946	
60	304155	H68696	
	304203	N56929	
	304234	W81608	
	304348	AA179868	
	304430	AA347682	
65	304456	AA411240	
	304521	AA464716	
	304526	AA476427	
	304607	AA513322	
	304735	AA576453	
70	304760	AA580401	
	306015	AA897116	
	306063	AA906316	
	306065	AA906725	
	306104	AA910956	
75	306109	AA911861	
	306242	AA932805	
	306288	AA936900	
	306396	AA970223	
	330568	NOT_FOUND_entrez	U56244
80	330599	15323_-12 U90437	
	331131	genbank_R54797	R54797
	331203	NOT_FOUND_entrez	T82310
	331531	genbank_N51343	N51343
	331547	467396_1 AA828597 N54811	
85	332074	genbank_AA599012	AA599012

TABLE 8C shows the genomic position for those Pkays in Table 8A lacking unigene ID's and accession numbers. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

5	Pkey:	Unique number corresponding to an Eos probe set		
	Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.		
	Strand:	Indicates DNA strand from which exons were predicted.		
10	Nt_position:	Indicates nucleotide positions of predicted exons.		
	Pkey	Ref	Strand	Nt_position
	332792	Dunham, I. et al.	Plus	73381-73768
	332816	Dunham, I. et al.	Plus	359844-360030
15	332906	Dunham, I. et al.	Plus	1923101-1923205
	332911	Dunham, I. et al.	Plus	1961767-1961858
	332912	Dunham, I. et al.	Plus	1962120-1962246
	332922	Dunham, I. et al.	Plus	2009620-2009738
	332956	Dunham, I. et al.	Plus	2510528-2510658
20	332959	Dunham, I. et al.	Plus	2518145-2518213
	333138	Dunham, I. et al.	Plus	3369205-3369323
	333139	Dunham, I. et al.	Plus	3369495-3369571
	333221	Dunham, I. et al.	Plus	3978070-3978187
	333380	Dunham, I. et al.	Plus	4904775-4904846
25	333387	Dunham, I. et al.	Plus	4910935-4910997
	333512	Dunham, I. et al.	Plus	5560510-5560564
	333524	Dunham, I. et al.	Plus	5612620-5612780
	333585	Dunham, I. et al.	Plus	6234778-6234894
	333618	Dunham, I. et al.	Plus	6562391-6562566
30	333627	Dunham, I. et al.	Plus	6620584-6620903
	333628	Dunham, I. et al.	Plus	6629004-6629233
	333650	Dunham, I. et al.	Plus	6796852-6797128
	333678	Dunham, I. et al.	Plus	7068223-7068288
	333750	Dunham, I. et al.	Plus	7608165-7608234
35	333763	Dunham, I. et al.	Plus	7692491-7692630
	333767	Dunham, I. et al.	Plus	7694407-7694623
	333768	Dunham, I. et al.	Plus	7695440-7695697
	333769	Dunham, I. et al.	Plus	7696625-7696707
	333772	Dunham, I. et al.	Plus	7706773-7706902
40	333777	Dunham, I. et al.	Plus	7746805-7746916
	333846	Dunham, I. et al.	Plus	8008623-8008757
	333884	Dunham, I. et al.	Plus	8153960-8154161
	333887	Dunham, I. et al.	Plus	8154892-8155025
45	333891	Dunham, I. et al.	Plus	8156437-8156709
	333892	Dunham, I. et al.	Plus	8156825-8157001
	333948	Dunham, I. et al.	Plus	8583497-8583627
	333954	Dunham, I. et al.	Plus	8563186-8563335
	333966	Dunham, I. et al.	Plus	8655643-8655826
50	333968	Dunham, I. et al.	Plus	8681004-8681241
	334061	Dunham, I. et al.	Plus	9686941-9687077
	334094	Dunham, I. et al.	Plus	9889953-9890105
	334113	Dunham, I. et al.	Plus	10282459-10282597
	334161	Dunham, I. et al.	Plus	10599033-10599180
55	334219	Dunham, I. et al.	Plus	12716160-12716384
	334239	Dunham, I. et al.	Plus	13056569-13056693
	334333	Dunham, I. et al.	Plus	13503544-13503657
	334378	Dunham, I. et al.	Plus	13907239-13907370
	334382	Dunham, I. et al.	Plus	13915866-13916036
60	334562	Dunham, I. et al.	Plus	14987847-14987940
	334588	Dunham, I. et al.	Plus	15032740-15032817
	334616	Dunham, I. et al.	Plus	15176123-15176470
	334633	Dunham, I. et al.	Plus	15333206-15333305
	334866	Dunham, I. et al.	Plus	18872214-18872317
65	334891	Dunham, I. et al.	Plus	19299770-19299944
	334934	Dunham, I. et al.	Plus	20103970-20104058
	335015	Dunham, I. et al.	Plus	20682792-20682945
	335120	Dunham, I. et al.	Plus	21436286-21436384
	335125	Dunham, I. et al.	Plus	21441390-21441471
	335179	Dunham, I. et al.	Plus	21634405-21634526
70	335188	Dunham, I. et al.	Plus	21669118-21669328
	335211	Dunham, I. et al.	Plus	21774611-21774680
	335361	Dunham, I. et al.	Plus	22807292-22807445
	335379	Dunham, I. et al.	Plus	22899306-22899420
	335414	Dunham, I. et al.	Plus	23235546-23235684
75	335416	Dunham, I. et al.	Plus	23237354-23237465
	335496	Dunham, I. et al.	Plus	24164366-24164545
	335497	Dunham, I. et al.	Plus	24167666-24167869
	335558	Dunham, I. et al.	Plus	24740167-24740347
	335586	Dunham, I. et al.	Plus	24990333-24990497
80	335686	Dunham, I. et al.	Plus	25439839-25439920
	335784	Dunham, I. et al.	Plus	25942710-25942792
	335823	Dunham, I. et al.	Plus	26365925-26366004
	335983	Dunham, I. et al.	Plus	27938968-27939070
	335995	Dunham, I. et al.	Plus	28009044-28009184
85	336021	Dunham, I. et al.	Plus	28686482-28686559

WO 02/086443

PCT/US02/12476

	336034	Dunham, I. et al.	Plus	29014404-29014590
	336038	Dunham, I. et al.	Plus	29022963-29023165
	336107	Dunham, I. et al.	Plus	29987731-29987859
5	336532	Dunham, I. et al.	Plus	983890-985529
	336533	Dunham, I. et al.	Plus	985591-986221
	336634	Dunham, I. et al.	Plus	986296-986570
	336635	Dunham, I. et al.	Plus	987908-988364
	336636	Dunham, I. et al.	Plus	988418-989185
10	336637	Dunham, I. et al.	Plus	989276-990813
	336638	Dunham, I. et al.	Plus	991906-993240
	336659	Dunham, I. et al.	Plus	1896402-1895478
	336694	Dunham, I. et al.	Plus	2420546-2420616
	336721	Dunham, I. et al.	Plus	3371522-3371586
	336900	Dunham, I. et al.	Plus	10236423-10236523
15	336948	Dunham, I. et al.	Plus	12692290-12692381
	337028	Dunham, I. et al.	Plus	16644817-16644942
	337054	Dunham, I. et al.	Plus	17821742-17821922
	337162	Dunham, I. et al.	Plus	23478943-23479145
	337183	Dunham, I. et al.	Plus	23943606-23943696
20	337184	Dunham, I. et al.	Plus	23973949-23974016
	337268	Dunham, I. et al.	Plus	28011979-28012034
	337299	Dunham, I. et al.	Plus	29022656-29022775
	337389	Dunham, I. et al.	Plus	31401509-31401579
	337493	Dunham, I. et al.	Plus	33330760-33330981
25	337549	Dunham, I. et al.	Plus	34474472-34474531
	337755	Dunham, I. et al.	Plus	3971764-3971900
	337809	Dunham, I. et al.	Plus	4449069-4449193
	337871	Dunham, I. et al.	Plus	5443027-5443101
30	337958	Dunham, I. et al.	Plus	6969162-6969270
	338008	Dunham, I. et al.	Plus	7697068-7697236
	338033	Dunham, I. et al.	Plus	8092128-8092271
	338110	Dunham, I. et al.	Plus	10384481-10384521
	338112	Dunham, I. et al.	Plus	10391398-10391600
35	338145	Dunham, I. et al.	Plus	11386629-11386692
	338148	Dunham, I. et al.	Plus	11448885-11449085
	338179	Dunham, I. et al.	Plus	12808775-12808833
	338197	Dunham, I. et al.	Plus	13638107-13638181
	338279	Dunham, I. et al.	Plus	16168944-16169091
40	338316	Dunham, I. et al.	Plus	17089711-17089988
	338322	Dunham, I. et al.	Plus	17132477-17132547
	338357	Dunham, I. et al.	Plus	18062184-18062402
	338359	Dunham, I. et al.	Plus	18074402-18074501
	338366	Dunham, I. et al.	Plus	18252026-18252189
45	338374	Dunham, I. et al.	Plus	18371200-18371282
	338414	Dunham, I. et al.	Plus	19345573-19345660
	338418	Dunham, I. et al.	Plus	19435506-19435596
	338501	Dunham, I. et al.	Plus	21244713-21244828
	338506	Dunham, I. et al.	Plus	21221871-21221953
50	338523	Dunham, I. et al.	Plus	21509763-21509864
	338662	Dunham, I. et al.	Plus	24404720-24404899
	338804	Dunham, I. et al.	Plus	27236005-27236108
	338836	Dunham, I. et al.	Plus	27792166-27792272
	338879	Dunham, I. et al.	Plus	28410553-28410734
55	338937	Dunham, I. et al.	Plus	29160555-29160725
	338993	Dunham, I. et al.	Plus	30077787-30078184
	339047	Dunham, I. et al.	Plus	30760793-30760968
	339100	Dunham, I. et al.	Plus	31141580-31141765
	339114	Dunham, I. et al.	Plus	31456454-31456519
60	339121	Dunham, I. et al.	Plus	31583467-31583536
	339170	Dunham, I. et al.	Plus	32216399-32216527
	339293	Dunham, I. et al.	Plus	33223671-33223819
	332858	Dunham, I. et al.	Minus	1339607-1339397
	332982	Dunham, I. et al.	Minus	2628296-2628109
65	332984	Dunham, I. et al.	Minus	2632606-2632457
	332998	Dunham, I. et al.	Minus	2711704-2711565
	333058	Dunham, I. et al.	Minus	3028925-3028811
	333097	Dunham, I. et al.	Minus	3204124-3204036
	333121	Dunham, I. et al.	Minus	3308446-3308358
70	333122	Dunham, I. et al.	Minus	3309596-3309531
	333123	Dunham, I. et al.	Minus	3310817-3310749
	333140	Dunham, I. et al.	Minus	3377220-3376309
	333260	Dunham, I. et al.	Minus	4308400-4308304
	333603	Dunham, I. et al.	Minus	6466335-6465727
	333604	Dunham, I. et al.	Minus	6467090-6466768
75	333904	Dunham, I. et al.	Minus	8217374-8217261
	333906	Dunham, I. et al.	Minus	8218238-8218063
	334183	Dunham, I. et al.	Minus	11832582-11832508
	334187	Dunham, I. et al.	Minus	11921456-11921205
80	334222	Dunham, I. et al.	Minus	12732417-12732289
	334223	Dunham, I. et al.	Minus	12734365-12734269
	334255	Dunham, I. et al.	Minus	13200776-13200692
	334492	Dunham, I. et al.	Minus	14478333-14478172
	334648	Dunham, I. et al.	Minus	15363301-15363222
	334787	Dunham, I. et al.	Minus	16299093-16298937
85	334933	Dunham, I. et al.	Minus	20078117-20077991

WO 02/086443

PCT/US02/12476

	334945	Dunham, I. et al.	Minus	20138885-20138637
	334967	Dunham, I. et al.	Minus	20173311-20173218
	334990	Dunham, I. et al.	Minus	20341159-20341087
5	335093	Dunham, I. et al.	Minus	21297367-21297214
	335288	Dunham, I. et al.	Minus	22304275-22303770
	335289	Dunham, I. et al.	Minus	22305950-22305708
	335548	Dunham, I. et al.	Minus	24662773-24662673
	335551	Dunham, I. et al.	Minus	24679828-24678961
10	335619	Dunham, I. et al.	Minus	25082677-25082498
	335620	Dunham, I. et al.	Minus	25092561-25092434
	335621	Dunham, I. et al.	Minus	25098878-25098767
	335682	Dunham, I. et al.	Minus	25421215-25421093
	335755	Dunham, I. et al.	Minus	25763805-25763747
	335814	Dunham, I. et al.	Minus	26320043-26319845
15	335815	Dunham, I. et al.	Minus	26320518-26320421
	335835	Dunham, I. et al.	Minus	26393311-26393245
	335851	Dunham, I. et al.	Minus	26604863-26604742
	335868	Dunham, I. et al.	Minus	26711437-26711300
20	335896	Dunham, I. et al.	Minus	26977639-26977558
	335936	Dunham, I. et al.	Minus	27360474-27360400
	335948	Dunham, I. et al.	Minus	27555924-27555788
	336066	Dunham, I. et al.	Minus	29241080-29240842
	336205	Dunham, I. et al.	Minus	30477456-30477311
25	336275	Dunham, I. et al.	Minus	32086675-32086536
	336292	Dunham, I. et al.	Minus	32818035-32817927
	336331	Dunham, I. et al.	Minus	33594527-33594371
	336419	Dunham, I. et al.	Minus	34052568-34052445
	336675	Dunham, I. et al.	Minus	2020758-2020664
30	336684	Dunham, I. et al.	Minus	2159060-2157993
	336716	Dunham, I. et al.	Minus	3259952-3259862
	336798	Dunham, I. et al.	Minus	5888954-5888757
	337043	Dunham, I. et al.	Minus	17407330-17407251
	337046	Dunham, I. et al.	Minus	17610892-17610821
35	337128	Dunham, I. et al.	Minus	22215251-22215034
	337192	Dunham, I. et al.	Minus	24591853-24591771
	337194	Dunham, I. et al.	Minus	24610510-24610359
	337229	Dunham, I. et al.	Minus	26716579-26716481
	337325	Dunham, I. et al.	Minus	30015948-30015800
40	337497	Dunham, I. et al.	Minus	33371317-33371258
	337500	Dunham, I. et al.	Minus	33376212-33376158
	337603	Dunham, I. et al.	Minus	1299296-1299194
	337605	Dunham, I. et al.	Minus	1346555-1346397
	337671	Dunham, I. et al.	Minus	3260634-3260547
45	337786	Dunham, I. et al.	Minus	4133203-4133081
	337862	Dunham, I. et al.	Minus	5347658-5347550
	338083	Dunham, I. et al.	Minus	9318438-9318301
	338158	Dunham, I. et al.	Minus	11794465-11794343
	338161	Dunham, I. et al.	Minus	12124716-12124658
50	338182	Dunham, I. et al.	Minus	12824919-12824827
	338189	Dunham, I. et al.	Minus	12878594-12878478
	338199	Dunham, I. et al.	Minus	13760865-13760780
	338215	Dunham, I. et al.	Minus	14055447-14055355
	338469	Dunham, I. et al.	Minus	20520387-20520242
55	338549	Dunham, I. et al.	Minus	22049171-22049081
	338561	Dunham, I. et al.	Minus	22311968-22311856
	338671	Dunham, I. et al.	Minus	24508421-24508346
	338676	Dunham, I. et al.	Minus	24637427-24637369
	338726	Dunham, I. et al.	Minus	25926206-25925618
50	338779	Dunham, I. et al.	Minus	27030151-27029795
	338871	Dunham, I. et al.	Minus	28301708-28301611
	338872	Dunham, I. et al.	Minus	28300921-28300790
	338966	Dunham, I. et al.	Minus	29614876-29614749
	339229	Dunham, I. et al.	Minus	32722330-32722199
55	339264	Dunham, I. et al.	Minus	32975145-32975053
	325228	6381940 Plus		2630-2694
	325235	6381943 Minus		162154-162264
	329588	3962484 Plus		1169-1619
	329560	3962491 Plus		2095-2990
70	329541	3983503 Minus		2765-3059
	325328	5866875 Plus		86780-86854
	325340	6017033 Minus		166656-166819
	325373	5866920 Minus		1136686-1136777
	325367	5866920 Minus		922881-922958
	325369	5866921 Plus		239672-239759
75	325436	5866939 Minus		29778-29907
	325498	5866967 Plus		173372-173930
	325471	6017034 Minus		289268-289342
	325557	6056302 Plus		50921-51050
30	325559	6249595 Minus		118590-119172
	325560	6249595 Minus		133794-133981
	325569	6249599 Plus		79927-80217
	325587	6682462 Plus		126724-126967
	325585	6682462 Plus		73476-73574
	325597	5866992 Plus		1065020-1065089
35	325639	5867002 Plus		253525-253608

WO 02/086443

	325739	5867038	Minus	205138-205269
	325740	5867038	Minus	207533-207690
	325792	6469828	Minus	1018-1176
5	325735	6552447	Minus	269122-269190
	325685	6682468	Plus	117397-117483
	325686	6682468	Plus	118337-118439
	325819	6682490	Minus	130314-130370
	325764	6048195	Minus	109733-109958
10	325703	6065793	Minus	139994-140138
	325643	6448539	Plus	53403-53537
	329816	6624888	Minus	70296-70423
	329860	6687260	Minus	163474-163605
	325883	5867087	Plus	22498-22663
15	325895	5867097	Plus	358317-358476
	325925	5867124	Plus	115749-115962
	325932	5867127	Plus	7369-7441
	325941	5867133	Minus	64228-64402
	325969	5867153	Plus	101911-102081
20	325971	5867153	Plus	105841-106035
	329993	4567166	Minus	101307-101434
	330020	6671887	Plus	172397-172491
	326163	5867168	Minus	7831-8035
	326274	5867171	Minus	410289-410404
25	326025	5867176	Plus	70854-70915
	326046	5867182	Minus	62668-62825
	326099	5867185	Minus	661381-661510
	326108	5867187	Minus	23784-23903
	326165	5867208	Minus	62787-62929
30	326189	5867212	Plus	69288-69413
	326204	5867218	Minus	148088-148200
	326230	5867230	Minus	301868-301972
	330052	4567182	Plus	352560-352963
	330036	6042048	Plus	117120-117216
35	326360	5867293	Plus	13627-13844
	326589	5867320	Plus	22760-22919
	326393	5867341	Plus	41702-41841
	326505	5867435	Minus	8818-8949
40	326515	5867439	Plus	36683-36809
	326592	6138928	Plus	23689-23828
	330107	6015249	Minus	100091-100282
	330106	6015249	Minus	99443-99778
	330100	6015253	Plus	21166-21301
	330093	6015278	Plus	1043-1199
45	330088	6015293	Plus	37517-37638
	330085	6015302	Minus	59613-59770
	330120	6671864	Minus	127553-127656
	330123	6671869	Minus	35311-35406
	326742	5867611	Minus	95187-95248
50	326605	5867637	Plus	24656-24749
	326818	6117831	Minus	15199-15309
	326720	6552456	Plus	84525-84677
	326770	6598307	Minus	513603-513668
	326692	6682502	Plus	117697-117899
55	326693	6682502	Minus	335002-335095
	326883	5867657	Minus	16023-16581
	326991	5867660	Plus	18147-18339
	326936	6004446	Minus	10217-10357
	326964	6469836	Plus	75340-75456
60	327040	6531965	Plus	783670-783817
	327053	6531965	Plus	2247267-2247437
	327075	6531965	Plus	4041318-4041431
	327085	6531965	Plus	4734947-4735069
	327036	6531965	Plus	319951-320040
	327130	6531976	Plus	20247-22343
65	327156	5866841	Minus	2462-2620
	327288	5867481	Plus	48583-48773
	327332	5867516	Minus	56361-56532
	327220	5867525	Minus	65701-65781
70	327224	5867534	Plus	188468-188544
	327321	6249562	Minus	99745-99836
	327361	6552412	Minus	61013-62130
	327396	5867743	Plus	8702-8820
	327414	5867750	Plus	102461-102586
75	327442	5867759	Plus	111483-111618
	327467	5867772	Plus	88030-88151
	327473	5867775	Plus	75101-75181
	327483	5867783	Plus	181573-181662
	327377	5867793	Minus	37610-37676
80	327562	5867804	Minus	343989-344474
	327568	5867811	Minus	46152-46287
	327606	6004463	Plus	200262-200495
	327611	5867868	Minus	175063-175392
	327642	5867891	Minus	2513-2743
85	327654	5867910	Minus	97564-97710
	327734	5867940	Minus	31003-31583

WO 02/086443

PCT/US02/12476

	327775	5867964	Minus	130791-130871
	327795	5867982	Plus	85267-85405
	327840	6249578	Minus	73065-73206
5	330208	6013599	Plus	66517-66931
	330263	6671884	Minus	101503-101634
	328004	5867993	Minus	157407-157887
	328101	5868020	Plus	289920-290014
	328100	5868020	Minus	263545-263635
10	328113	5868024	Minus	80378-80491
	328157	5868064	Plus	73326-73615
	328195	5868080	Minus	16551-16729
	328197	5868081	Minus	42133-42438
	327940	5868197	Minus	95240-95428
15	327984	5868216	Plus	66611-66677
	328021	5902482	Plus	713478-714590
	328068	6117819	Plus	253903-254022
	328264	6381912	Plus	55086-55404
	330300	2905862	Minus	3246-3302
20	328608	5868222	Minus	87770-87953
	328600	5868229	Minus	38889-40010
	328616	5868239	Plus	293920-294224
	328623	5868246	Minus	120020-120126
	328632	5868247	Plus	76734-76853
	328666	5868254	Minus	778-901
25	328698	5868264	Minus	625555-625633
	328700	5868264	Plus	764089-764203
	328708	5868271	Minus	68114-68854
	328735	5868289	Plus	89389-89455
30	328743	5868289	Plus	274638-274726
	328806	5868324	Plus	29408-29684
	328299	5868366	Minus	149708-149889
	328342	5868383	Plus	59955-60094
	328365	5868387	Minus	270724-270798
35	328369	5868388	Plus	75371-75583
	328381	5868392	Plus	662758-662848
	328451	5868425	Minus	217275-217336
	328481	5868449	Minus	8987-9180
	328500	5868464	Plus	59098-59481
40	328530	5868482	Plus	334973-335406
	328664	6004473	Plus	1193739-1193866
	328861	6381928	Minus	108317-108403
	328908	5868493	Plus	117002-117059
	328933	5868500	Plus	771755-771889
45	328934	5868500	Plus	846342-846448
	328949	6456765	Minus	43552-43619
	330313	6042030	Minus	33642-33775
	329005	5868542	Plus	85470-85673
	330366	2944106	Plus	151837-151914
50	330372	6580495	Minus	317461-317688
	329033	5868561	Minus	5390-5479
	329037	5868562	Minus	32466-32562
	329067	5868591	Minus	146417-147652
	329134	5868679	Plus	29969-30018
55	329157	5868687	Minus	145940-146155
	329178	5868704	Plus	179177-179463
	329192	5868716	Plus	166936-167020
	329194	5868716	Minus	304450-304559
	329204	5868720	Minus	3050-3190
60	329224	5868728	Plus	27422-27664
	329228	5868728	Minus	50118-50287
	329288	5868771	Plus	25554-26299
	329337	5868806	Minus	467155-467222
	329011	6682532	Plus	48658-48741

TABLE 9A: Potential Therapeutic, Diagnostic and Prognostic targets for Therapy of Lung Cancer

Table 9A shows about 1312 genes up-regulated in lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) relative to normal body tissues. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 9B show the accession numbers for those Pkey's lacking UnigenelD's for table 9A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 9C show the genomic positioning for those Pkey's lacking UnigenelD's and accession numbers in table 9A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number
 ExAccon: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigenel number
 Unigenel Title: Unigenel gene title
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccon	UnigenelD	Unigenel Title	R1	R2
400195			NM_007057*:Homo sapiens ZW10 interactor	1.00	1.00
400205			NM_006265*:Homo sapiens RAD21 (S. pombe)	15.80	396.00
400220			Eos Control	2.28	2.84
400277			Eos Control	7.68	9.72
400285			Eos Control	1.00	1.00
400288	X06256	Hs.149609	integrin, alpha 5 (fibronectin receptor,	1.04	2.24
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	132.45	4.00
400298	AA032279	Hs.61635	six transmembrane epithelial antigen of	43.86	74.00
400301	X03635	Hs.1657	estrogen receptor 1	1.00	1.00
400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	1.75	1.65
400328	X87344	Hs.180062	transporter 2, ATP-binding cassette, sub	0.87	1.80
400419	AF084545		Target	156.55	253.00
400512			NM_030878*:Homo sapiens cytochrome P450,	1.00	2.00
400517	AF242388		lengsin	3.57	87.00
400560			NM_030878*:Homo sapiens cytochrome P450,	1.00	1.00
400664			NM_002425*:Homo sapiens matrix metallopro	20.26	45.00
400665			NM_002425*:Homo sapiens matrix metallopro	1.36	1.07
400666			NM_002425*:Homo sapiens matrix metallopro	3.26	3.22
400749			NM_003105*:Homo sapiens sortilin-related	1.00	91.00
400763			Target Exon	7.63	24.00
401027			Target Exon	1.00	1.00
401093			C12000586*:gi 6330167 dbj BAA86477.1 (A	1.00	155.00
401203			Target Exon	1.00	86.00
401212			C12000457*:gi 7512178 pir T30337 polypr	1.00	400.00
401411			ENSP00000247172*:HYPOTHETICAL 126.2 kDa	1.00	72.00
401435			C14000397*:gi 7499898 pir T33295 hypoth	1.00	64.00
401464	AF039241		histone deacetylase 5	3.82	49.00
401714			ENSP00000241802*:CDNA FLJ11007 FIS, CLON	2.02	40.00
401747			Homo sapiens keratin 17 (KRT17)	128.43	68.00
401760			Target Exon	1.74	35.00
401780			NM_005557*:Homo sapiens keratin 16 (foca	26.47	10.50
401781			Target Exon	10.33	4.61
401785			NM_002275*:Homo sapiens keratin 15 (KRT1	4.13	2.70
401797			Target Exon	1.44	2.10
401961			NM_012626*:Homo sapiens serine carboxypep	1.41	1.86
401985	AF053004		class I cytokine receptor	1.00	177.00
401994			Target Exon	61.84	47.00
402075			ENSP00000251056*:Plasma membrane calcium	1.00	1.00
402260			NM_001436*:Homo sapiens fibrillarin (FBL	1.58	1.39
402265			Target Exon	2.09	35.00
402297			Target Exon	1.00	92.00
402408			NM_030920*:Homo sapiens hypothetical pro	28.87	13.00
402420			C1000823*:gi 10432400 amb CAC10290.1 (A	1.00	1.44
402674			Target Exon	7.44	243.00
402802			NM_001397*:Homo sapiens endothenin conver	1.00	70.00
402994			NM_002463*:Homo sapiens myxovirus (influ	1.37	1.43
403137			NM_005381*:Homo sapiens nucleolin (NCL),	1.00	19.00
403306	NM_006825		transmembrane protein (53kD), endoplasmic	1.00	43.00
403329			Target Exon	1.00	61.00
403381			ENSP00000231844*:Ecotropic virus integra	1.00	119.00
403478			NM_022342*:Homo sapiens kinesin protein 9	28.13	136.00
403485			C3001813*:gi 12737279 ref XP_012163.1 k	20.23	76.00
403627			Target Exon	6.30	29.33
403715			Target Exon	1.30	35.00
404044			ENSP00000237855*:DJ399G3.2 (NOVEL PROTEI	1.00	54.00
404076			NM_016020*:Homo sapiens CGI-75 protein (14.29	91.00
404101			C8000950*:gi 423560 pir A47318 RNA-bindi	1.00	1.00
404140			NM_006510*:Homo sapiens ret finger protai	1.42	1.44
404165			ENSP00000244562*:NRH dehydrogenase [quino	1.00	54.00
404185			Target Exon	1.00	117.00
404210			NM_005936*:Homo sapiens myeloid/lymphoid	5.93	13.77
404253			NM_021058*:Homo sapiens H2B histone fami	1.00	1.00

404287			C6001909:gi 704441 kb BAA18909.1 [D298	29.71	42.00
404298			C6001238*:gi 121715 sp P25697 GTA3_CHICK	1.30	1.00
404347			Target Exon	1.00	1.00
404440			NM_021048:Homo sapiens melanoma antigen,	1.00	15.00
5 404721			NM_005596*:Homo sapiens nuclear factor I	1.00	60.00
404794	NM_000078		cholesterol ester transfer protein, plas	1.07	1.38
404854			Target Exon	1.61	2.01
404877			NM_005365:Homo sapiens melanoma antigen,	1.00	1.00
404927			Target Exon	1.00	1.00
10 404996			Target Exon	1.00	1.00
405449			CY000047*:gi 11427234 ref XP_009399.1 z	1.00	1.00
405568			NM_031413*:Homo sapiens cat eye syndrome	1.00	78.00
405572			Target Exon	0.76	1.14
405646			C12000200:gi 4557225 ref NP_000005.1 al	1.01	1.28
15 405676	BE336714		cytochrome c-1	1.13	2.89
405770			NM_002362:Homo sapiens melanoma antigen,	45.52	37.00
405932			C15000305:gi 3806122 gb AAC69198.1 (AF0	1.99	1.99
406137			NM_000179*:Homo sapiens mutS (E. coli) h	2.77	2.38
406360			Target Exon	1.00	35.00
20 406399			NM_003122*:Homo sapiens serine protease	1.00	39.00
406467			Target Exon	1.00	1.00
406621	X57809	Hs.181125	immunoglobulin lambda locus	1.41	1.74
406642	AJ245210		gb:Homo sapiens mRNA for immunoglobulin	2.16	3.91
406663	U24683	Hs.293441	immunoglobulin heavy constant mu	2.07	2.93
25 406671	AA129547	Hs.285754	met proto-oncogene (hepatocyte growth fa	15.00	51.00
406673	M34996	Hs.198253	major histocompatibility complex, class	0.98	3.09
406676	X58399	Hs.81221	Human L2-9 transcript of unrearranged Im,	1.30	1.53
406678	U77534		gb:Human clone 1A11 immunoglobulin varia	1.33	1.45
406685	M18728		gb:Human nonspecific crossreacting anti	1.46	2.85
30 406687	M31126	Hs.272822	pregnancy specific beta-1-glycoprotein 9	8.61	8.50
406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad	226.37	350.00
406698	X03068	Hs.73931	major histocompatibility complex, class	1.01	2.52
406815	AA833930	Hs.288036	tRNA isopentanylpuriphosphate transferas	20.25	32.00
406851	AA609784		major histocompatibility complex, class	0.75	1.91
35 406964	M21305		gb:Human alpha satellite and satellite 3	38.15	1114.00
406967	M24349		gb:Human parathyroid hormone-like protei	1.00	1.00
406974	M57293		gb:Human parathyroid hormone-related pep	1.00	1.00
407103	AA424881	Hs.256301	hypothetical protein MGC13170	1.77	1.10
407128	R83312	Hs.237260	EST	1.00	1.00
40 407137	T97307		gb:ye53h05.s1 Soares fetal liver spleen	142.70	135.00
407168	R45175	Hs.117183	ESTs	2.16	18.00
407239	AA076350	Hs.67846	leukocyte immunoglobulin-like receptor,	1.10	1.57
407242	M18728		gb:Human nonspecific crossreacting anti	1.12	2.85
407244	M10014	Hs.75431	fibrinogen, gamma polypeptide	3.24	15.38
45 407289	AA135159	Hs.203349	Homo sapiens cDNA FLJ12149 fis, clone MA	3.53	3.68
407300	AA102616	Hs.120769	gb:zn43e07.s1 Stratagene HeLa cell s3 93	19.74	73.00
407366	AF026942	Hs.271530	gb:Homo sapiens cig33 mRNA, partial sequ	0.06	8.25
407378	AA299264	Hs.57776	ESTs, Moderately similar to I38022 hypot	1.00	26.00
407430	AF169351		gb:Homo sapiens protein tyrosine phospho	1.00	25.00
50 407453	AJ132087		gb:Homo sapiens mRNA for axonemal dynein	1.00	75.00
407577	AW131324	Hs.246759	hypothetical protein MGC12538	1.00	1.00
407634	AW016569	Hs.136414	UDP-GlcNAc:betaGal beta-1,3-N-acetylgluc	111.20	228.00
407710	AW022727	Hs.23616	ESTs	1.00	28.00
407720	AB037776	Hs.38002	KIAA1355 protein	1.89	1.31
55 407746	AK001962		hypothetical protein FLJ11100	1.00	1.00
407756	AA116021	Hs.38260	ubiquitin specific protease 18	4.51	5.00
407758	D50915	Hs.38365	KIAA0125 gene product	1.00	28.00
407782	AA608956	Hs.112619	ESTs, Moderately similar to PURKINJE CEL	0.97	1.14
407788	BE514982	Hs.38991	S100 calcium-binding protein A2	7.88	3.83
60 407790	AJ027274	Hs.288941	Homo sapiens cDNA FLJ14866 fis, clone PL	3.63	42.00
407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	89.96	109.00
407839	AA045144	Hs.161566	ESTs	173.91	108.00
407944	R34008	Hs.239727	desmocollin 2	111.30	70.00
408000	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	151.17	8.00
65 408031	AA081395	Hs.42173	Homo sapiens cDNA FLJ10366 fis, clone NT	9.91	93.00
408063	BE085548	Hs.42346	calcineurin-binding protein calcisarcin-1	195.78	231.00
408070	AW148852		gb:xf05d05.x1 NCL_CGAP_Bm35 Homo sapien	1.00	1.00
408101	AW988504	Hs.123073	CDC2-related protein kinase 7	37.84	61.00
408122	AJ432652	Hs.42824	hypothetical protein FLJ10718	0.85	1.71
70 408212	AA297567	Hs.43728	hypothetical protein	5.88	7.91
408243	Y00787	Hs.624	interleukin 8	4.27	9.98
408349	BE546947	Hs.44276	homeo box C10	3.79	3.46
408353	BE439838	Hs.44298	mitochondrial ribosomal protein S17	1.88	1.65
408354	AJ382803	Hs.159235	ESTs	1.00	73.00
75 408369	R38438	Hs.182575	solute carrier family 15 (H777) transport	1.41	16.50
408380	AF123050	Hs.44532	diubiquitin	15.19	37.22
408482	NM_000676	Hs.45743	adenosine A2b receptor	1.65	1.19
408522	AJ541214	Hs.46320	Small proline-rich protein SPRK (human,	1.98	1.24
408536	AW381532	Hs.135188	ESTs	1.55	1.50
80 408545	AW235405	Hs.253690	ESTs	1.00	1.00
408572	AA056511	Hs.226568	ESTs, Moderately similar to ALU4_HUMAN A	1.00	44.00
408633	AW963372	Hs.46677	PRO2000 protein	107.16	56.00
408660	AA525775		ESTs, Moderately similar to PC4259 ferri	1.00	1.00
408761	AA057264	Hs.238936	ESTs, Weakly similar to (define not ava	52.24	141.00
85 408771	AW732573	Hs.47584	potassium voltage-gated channel, delayed	3.05	109.00

	408783	AF192522	Hs.47701	NPC1 (Niemann-Pick disease, type C1, gen	1.02	1.07
	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	41.19	61.00
	408805	H69912	Hs.48269	vaccinia related kinase 1	24.67	45.00
5	408841	AW438865	Hs.256862	ESTs	1.00	58.00
	408873	AL046017	Hs.182278	calmodulin 2 (phosphorylase kinase, deli	1.00	89.00
	408908	BE296227	Hs.250822	serine/threonine kinase 15	7.76	1.00
	408992	AA059325	Hs.71642	guanine nucleotide binding protein (G pr	1.00	1.00
	408996	AI979168	Hs.344096	glycoprotein (transmembrane) numb	3.71	5.50
10	409015	BE389387	Hs.49767	NM_004553:Homo sapiens NADH dehydrogenas	1.44	1.24
	409038	T97490	Hs.50002	small inducible cytokine subfamily A (Cy	4.28	5.32
	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	112.42	195.00
	409077	AA401369	Hs.190721	ESTs	1.00	17.00
	409093	BE243834	Hs.50441	CGI-04 protein	2.02	1.93
15	409103	AF251237	Hs.112208	XAGE-1 protein	80.44	40.00
	409142	AL136877	Hs.50758	SMC4 (structural maintenance of chromoso	14.87	6.00
	409187	AF154830	Hs.50966	carbamoyl-phosphate synthetase 1, mitoch	1.00	1.00
	409228	AI654298	Hs.271695	ESTs, Weakly similar to 2109260A B cell	1.22	1.00
	409234	AI879419	Hs.27206	ESTs	1.00	1.00
20	409268	AA625304	Hs.187579	ESTs	11.90	23.00
	409269	AA576953	Hs.22972	hypothetical protein FLJ13352	1.00	1.00
	409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	168.91	35.00
	409404	BE220053	Hs.129056	ESTs	1.00	1.00
	409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini	79.74	96.00
25	409430	R21945	Hs.346735	splicing factor, arginine/serine-rich 5	1.45	2.10
	409446	AI561173	Hs.67688	ESTs	1.00	4.00
	409506	NM_006153	Hs.54589	NCK adaptor protein 1	3.97	28.00
	409522	AA075382		gb:zm87b03.s1 Stratagene ovarian cancer	15.98	141.00
	409582	AA401369	Hs.190721	ESTs	1.00	17.00
30	409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibito	292.12	79.00
	409705	M37762	Hs.56023	brain-derived neurotrophic factor	1.00	82.00
	409719	AI769160	Hs.108681	Homo sapiens brain tumor associated prot	1.00	1.00
	409731	AA125985	Hs.56145	thymosin, beta, identified in neuroblast	0.12	18.12
	409744	AW675258	Hs.56265	Homo sapiens mRNA; cDNA DKFZp586P2321 (f	20.75	51.00
	409757	NM_001898	Hs.123114	cystatin SN	22.46	15.80
35	409866	AW502152		gb:U1-HF-BR0p-ajr-f-11-Q-U1r1 NIH_MGC_5	1.00	1.00
	409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.	1.50	1.09
	409902	AI337658	Hs.156351	ESTs	25.92	50.00
	409935	AW511413	Hs.278025	ESTs	2.63	2.11
40	409956	AW103364	Hs.727	inhibin, beta A (activin A, activin AB a	2.17	4.01
	409958	NM_001523	Hs.57697	hyaluronan synthase 1	0.91	2.07
	410001	AB041036	Hs.57771	kallikrein 11	1.04	2.28
	410032	BE065985		gb:RC3-BT0319-120200-014-a09 BT0319 Homo	1.00	58.00
	410037	AB020725	Hs.58009	KIAA0918 protein	1.00	34.00
45	410044	BE566742	Hs.58169	highly expressed in cancer, rich in leuc	1.00	1.00
	410048	W76467	Hs.58218	proline oxidase homolog	1.03	1.44
	410076	T05387	Hs.7991	ESTs	1.12	1.50
	410102	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	9.89	1.00
	410153	BE311926	Hs.15830	hypothetical protein FLJ12691	1.00	1.00
	410166	AK001376	Hs.59346	hypothetical protein FLJ10514	1.00	1.00
50	410193	AI132592	Hs.59757	zinc finger protein 281	42.01	51.00
	410274	AA381807	Hs.61762	hypoxia-inducible protein 2	1.72	1.32
	410309	BE043077	Hs.278153	ESTs	1.00	2.00
	410340	AW182833	Hs.112188	hypothetical protein FLJ13149	32.08	75.00
55	410348	AW182663	Hs.95469	ESTs	1.00	1.00
	410407	X66839	Hs.63287	carbonic anhydrase IX	1.40	1.11
	410418	D31382	Hs.63325	transmembrane protease, serine 4	4.30	2.03
	410438	AB037756	Hs.45207	hypothetical protein KIAA1335	1.00	18.00
	410553	AW016824	Hs.255527	hypothetical protein MGC14128	1.34	1.04
60	410555	W27235	Hs.64311	a disintegrin and metalloproteinase doma	23.99	1.41
	410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	10.04	1.00
	410681	AW246890	Hs.65425	calbindin 1, (28kD)	10.88	18.92
	410781	AI375672	Hs.165028	ESTs	1.00	57.00
	411027	AF072099	Hs.67846	leukocyte immunoglobulin-like receptor,	1.62	3.78
65	411074	X60435	Hs.68137	adenylate cyclase activating polypeptide	1.00	1.15
	411089	AA456454		cell division cycle 2-like 1 (PITSLRE pr	1.55	1.58
	411152	BE069199		gb:QV3-BT0379-010300-105-g03 BT0379 Homo	1.00	84.00
	411248	AA551538	Hs.334605	Homo sapiens cDNA FLJ14408 fis, clone HE	1.82	1.45
	411252	AB018549	Hs.69328	MD-2 protein	7.32	12.74
70	411263	BE297802	Hs.69360	kinesin-like 6 (mitotic centromere-assoc	3.44	2.55
	411365	M76477	Hs.289082	GM2 ganglioside activator protein	1.35	2.02
	411402	BE297855	Hs.69855	NRAS-related gene	1.00	46.00
	411573	AB029000	Hs.70823	KIAA1077 protein	11.40	11.35
	411579	AC005258	Hs.70830	U6 snRNA-associated Sm-like protein LSm7	1.08	1.90
75	411617	AA247994	Hs.50063	neurocalcin delta	1.74	2.57
	411732	AA059325	Hs.71642	guanine nucleotide binding protein (G pr	1.02	1.00
	411773	NM_006799	Hs.72026	protease, serine, 21 (testisin)	1.34	2.19
	411789	AF245505	Hs.72157	Adlican	2.19	2.79
	411800	N39342	Hs.103042	microtubule-associated protein 18	23.34	34.00
80	411945	AL033527	Hs.92137	v-myc avian myelocytomatosis viral oncog	1.00	8.00
	412115	AK001763	Hs.73239	hypothetical protein FLJ10901	2.07	1.64
	412140	AA219691	Hs.73625	RAB6 interacting, kinesin-like (raklines	118.48	92.00
	412276	BE262621	Hs.73798	macrophage migration inhibitory factor (1.58	1.49
	412464	T78141	Hs.22826	ESTs, Weakly similar to I55214 salivary	1.16	1.34
85	412530	AA766268	Hs.265273	hypothetical protein FLJ13346	41.52	84.00
	412537	AL031778		nuclear transcription factor Y, alpha	17.90	55.00

	412659	AW753855	Hs.74376	olfactomedin related ER localized protal	14.65	47.00
	412719	AW016510	Hs.816	ESTs	382.46	128.00
	412723	AA648459	Hs.335951	hypothetical protein AF301222	54.90	1.00
	412811	H06382		ESTs	1.00	11.00
5	412817	AL037159	Hs.74619	proteasome (prosome, macropain) 26S subu	1.63	1.42
	412863	AA121673	Hs.59757	zinc finger protein 281	17.63	56.00
	412924	BE018422	Hs.75258	H2A histone family, member Y	1.00	22.00
	413004	T35901	Hs.75117	interleukin enhancer binding factor 2, 4	2.19	2.05
10	413011	AW068115	Hs.821	biglycan	1.22	1.88
	413048	M93221	Hs.75182	mannose receptor, C type 1	0.30	6.23
	413063	AL035737	Hs.75184	chitinase 3-like 1 (cartilage glycoprote	3.43	8.71
	413129	AF292100	Hs.104613	RP42 homolog	4.67	4.77
	413142	M81740	Hs.75212	ornithine decarboxylase 1	1.92	2.59
	413223	AJ732182	Hs.191866	ESTs	5.73	27.00
15	413248	T64658	Hs.21433	hypothetical protein DKFZp547J035	0.99	1.06
	413273	U75579	Hs.75257	stem-loop (histone) binding protein	1.00	18.00
	413278	BE563085	Hs.833	interferon-stimulated protein, 15 kDa	1.10	1.09
	413281	AA861271	Hs.222024	transcription factor BMAL2	95.94	69.00
	413364	BE536218	Hs.137516	fidgetin-like 1	1.00	1.00
20	413385	M34455	Hs.840	indoleamine-pyrrole 2,3 dioxygenase	0.95	2.09
	413409	AI638418	Hs.1440	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	1.00	1.00
	413453	AA129640	Hs.128055	ESTs	1.00	31.00
	413527	BE250788	Hs.179882	hypothetical protein FLJ12443	1.08	1.46
	413554	AA319146	Hs.75426	secretogranin II (chromogranin C)	79.15	114.00
25	413573	AJ733859	Hs.149089	ESTs	1.00	1.00
	413582	AW295647	Hs.71331	hypothetical protein MGC5350	8.80	10.00
	413597	AW302885	Hs.117183	ESTs	1.00	1.00
	413690	BE157489		gb:RC1-HT0375-120200-011-e06 HT0375 Homo	1.00	1.00
30	413691	AB023173	Hs.75478	ATPase, Class VI, type 11B	3.16	2.32
	413719	BE439580	Hs.75498	small inducible cytokine subfamily A (Cy	2.88	9.52
	413753	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kafirin	144.10	108.00
	413801	M62246	Hs.35406	ESTs, Highly similar to unnamed protein	1.00	17.00
	413833	Z15005	Hs.75573	centromere protein E (312kD)	1.00	1.00
35	413882	AA132973	Hs.184492	ESTs	64.24	148.00
	413926	AA133338	Hs.54310	ESTs	1.00	67.00
	413943	AW294416	Hs.144687	Homo sapiens cDNA FLJ12981 fis, clone NT	43.42	42.00
	413995	BE048146	Hs.75571	syntaxin 1A (brain)	1.23	1.11
	414035	Y00630	Hs.75716	serine (or cysteine) proteinase inhibitor	2.02	2.51
40	414142	AW368397	Hs.334465	Homo sapiens cDNA FLJ14438 fis, clone HE	1.00	102.00
	414180	AI863304	Hs.120905	Homo sapiens cDNA FLJ11448 fis, clone HE	6.92	77.00
	414245	BE148072	Hs.75850	WAS protein family, member 1	1.00	1.00
	414275	AW970254	Hs.889	Charol-Leyden crystal protein	1.00	59.00
	414317	BE263280	Hs.75888	phosphogluconate dehydrogenase	1.52	1.73
45	414334	AA824298	Hs.21331	hypothetical protein FLJ10036	1.78	1.72
	414341	D80004	Hs.75909	KIAA0182 protein	33.90	151.00
	414368	W70171	Hs.75939	uridine monophosphate kinase	171.60	97.00
	414416	AW409985	Hs.76084	hypothetical protein MGC2721	2.32	1.85
	414430	AJ346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1	226.15	66.00
50	414570	Y00285	Hs.76473	insulin-like growth factor 2 receptor	1.64	1.98
	414618	AI204600	Hs.96978	hypothetical protein MGC10764	1.87	72.00
	414675	R79015	Hs.296281	Interleukin enhancer binding factor 1	1.51	1.39
	414683	S78295	Hs.76888	hypothetical protein MGC12702	43.61	64.00
	414696	AF002020	Hs.76918	Niemann-Pick disease, type C1	28.63	71.00
55	414711	AJ310440	Hs.288735	Homo sapiens cDNA FLJ13522 fis, clone PL	14.86	42.00
	414718	H95348	Hs.107987	ESTs	1.00	5.00
	414732	AW410976	Hs.77152	minichromosome maintenance deficient (S.	1.64	1.44
	414747	U30872	Hs.77204	centromere protein F (350/400kD, mitotin	65.01	74.00
	414761	AJ077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2	130.35	121.00
60	414774	X02419	Hs.77274	plasminogen activator, urokinase	2.24	2.19
	414806	D14694	Hs.77329	phosphatidylserine synthase 1	1.63	1.53
	414809	AJ434699	Hs.77356	transferrin receptor (p90, CD71)	1.97	2.60
	414812	X72755	Hs.77367	monokine induced by gamma interferon	3.48	10.60
	414825	X06370	Hs.77432	epidermal growth factor receptor (avian	103.22	143.00
	414839	X63692	Hs.77462	DNA (cytosine-5)-methyltransferase 1	1.80	1.69
65	414883	AA926960		CDC28 protein kinase 1	14.29	10.06
	414907	X90725	Hs.77597	polo (Drosophila)-like kinase	1.95	2.20
	414914	U49844	Hs.77613	ataxia telangiectasia and Rad3 related	3.00	2.90
	414945	BE076358	Hs.77667	lymphocyte antigen 6 complex, locus E	1.02	1.21
	414972	BE263782	Hs.77695	KIAA0008 gene product	1.00	1.00
70	415014	AW954064	Hs.24951	ESTs	1.42	2.84
	415091	AL044872	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A sy	1.00	30.00
	415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	34.72	107.00
	415227	AW821113	Hs.72402	ESTs	1.87	49.00
75	415238	R37780	Hs.21422	ESTs	1.00	1.00
	415263	AA948033	Hs.130853	ESTs	1.00	1.00
	415295	R41450	Hs.6546	ESTs	1.00	1.00
	415339	NM_015156	Hs.78398	KIAA0071 protein	51.18	166.00
	415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor	30.84	63.00
80	415674	BE394784	Hs.78596	proteasome (prosome, macropain) subunit,	1.48	1.39
	415709	AA649850	Hs.278558	ESTs	1.00	1.00
	415735	AA704162	Hs.120811	ESTs, Weakly similar to 138022 hypotheti	1.00	72.00
	415799	AA653718	Hs.225841	DKFZP434D193 protein	6.23	31.00
	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	24.30	1.00
85	415857	AA866115	Hs.127797	Homo sapiens cDNA FLJ11381 fis, clone HE	32.51	35.00
	415889	AI267700		ESTs	78.89	1.00

	416018	AW138239	Hs.78977	proprotein convertase subtilisin/kexin 1	1.00	1.00
	416065	BE267931	Hs.78996	proliferating cell nuclear antigen	3.35	2.32
	416111	AA033813	Hs.79018	chromatin assembly factor 1, subunit A (39.03	3.00
5	416177	AA174069	Hs.187607	ESTs	1.00	9.00
	416178	AI808527	Hs.192822	serologically defined breast cancer anti	3.83	3.76
	416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	3.67	1.00
	416209	AA235776	Hs.79078	MAD2 (mitotic arrest deficient, yeast, h	9.70	1.00
	416239	AL038450	Hs.48948	ESTs	83.87	129.00
10	416250	AA581386	Hs.73452	hypothetical protein MGC10791	1.96	2.12
	416322	BE019494	Hs.79217	pyrrolina-5-carboxylate reductase 1	2.08	1.73
	416423	H54375	Hs.268921	ESTs	1.00	89.00
	416448	L13210	Hs.79339	lectin, galactoside-binding, soluble, 3	1.28	1.54
	416498	U33632	Hs.79351	potassium channel, subfamily K, member 1	27.29	67.00
15	416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	53.29	51.00
	416661	AA634543	Hs.79440	IGF-II mRNA-binding protein 3	9.96	5.00
	416722	AA354604	Hs.122546	hypothetical protein FLJ23017	3.68	33.00
	416819	U77735	Hs.80205	pim-2 oncogene	1.59	1.84
	416836	N21352	Hs.42987	ESTs, Weakly similar to S21348 probable	1.00	1.00
20	417034	NM_006183	Hs.80962	neurotensin	1.00	1.00
	417061	AI675944	Hs.188691	Homo sapiens cDNA FLJ12033 fis, clone HE	32.95	156.00
	417079	U65590	Hs.81134	interleukin 1 receptor antagonist	3.91	4.93
	417218	AA129547	Hs.285754	mel proto-oncogene (hepatocyte growth fa	1.00	51.00
	417233	W25005	Hs.24395	small inducible cytokine subfamily B (Cy	3.38	2.05
25	417308	H60720	Hs.81892	KIAA0101 gene product	82.94	25.36
	417315	AI080042	Hs.180450	ribosomal protein S24	106.61	121.00
	417324	AW265494		ESTs	1.20	1.28
	417366	BE185289	Hs.1076	small proline-rich protein 1B (cornifin)	8.97	3.27
30	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor	2.59	1.82
	417428	N87579	Hs.278971	gb:LL2030F Human fetal heart, Lambda ZAP	1.00	52.00
	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	304.75	173.00
	417466	AI681547	Hs.59457	hypothetical protein FLJ22127	1.24	1.34
	417512	AI979168	Hs.344096	glycoprotein (transmembrane) nmb	2.14	5.50
	417515	L24203	Hs.82237	ataxia-telangiectasia group D-associated	2.66	1.68
35	417542	J04129	Hs.82269	progesterone-associated endometrial prote	1.28	1.35
	417576	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	42.76	51.00
	417715	AW969587	Hs.86366	ESTs	6.35	2.75
	417720	AA205625	Hs.208067	ESTs	113.31	56.00
	417791	AW965339	Hs.111471	ESTs	39.98	18.00
40	417830	AW504786	Hs.122579	hypothetical protein FLJ10461	2.61	31.00
	417866	AW067903	Hs.82772	collagen, type XI, alpha 1	2.35	2.44
	417900	BE250127	Hs.82906	CDC20 (cell division cycle 20, S. cerevi	1.52	1.11
	417933	X02308	Hs.82962	thymidylate synthetase	4.74	2.55
	417944	AU077196	Hs.82985	collagen, type V, alpha 2	3.61	5.21
45	417975	AA641836	Hs.30085	hypothetical protein FLJ23186	12.49	38.00
	417991	AA731452	Hs.190008	ESTs	1.00	26.00
	418004	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member	3.02	2.12
	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	187.59	1.00
	418054	NM_002318	Hs.83354	lysyl oxidase-like 2	2.85	2.63
50	418057	NM_012151	Hs.83363	coagulation factor VIII-associated (intr	1.54	1.69
	418113	AI272141	Hs.83484	SRY (sex determining region Y)-box 4	6.82	5.22
	418140	BE613836	Hs.83551	microfibrillar-associated protein 2	1.26	1.46
	418203	X54942	Hs.83758	CDC28 protein kinase 2	134.19	144.00
	418207	C14685	Hs.34772	ESTs	1.00	1.00
55	418216	AA662240	Hs.283099	AF15q14 protein	64.66	61.00
	418236	AW994005	Hs.337534	ESTs	18.53	147.00
	418249	H89226	Hs.34892	KIAA1323 protein	30.53	106.00
	418281	U09550	Hs.1154	oviductal glycoprotein 1, 120kD (mucin 9	1.00	3.00
	418283	S79895	Hs.83942	cathepsin K (pseudosynostosis)	3.96	5.16
60	418300	AI433074	Hs.86682	Homo sapiens cDNA: FLJ21578 fis, clone C	3.18	2.91
	418322	AA284166	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	11.96	6.68
	418327	U70370	Hs.84136	paired-like homeodomain transcription fa	9.23	2.22
	418345	AJ001696	Hs.241407	serine (or cysteine) proteinase inhibitor	1.00	1.00
	418379	AA218940	Hs.137516	fidgulin-like 1	21.68	44.00
65	418397	NM_001269	Hs.84746	chromosome condensation 1	1.00	8.00
	418403	D86978	Hs.84790	KIAA0225 protein	16.91	18.98
	418462	BE001596	Hs.85268	integrin, beta 4	1.56	1.16
	418478	U36945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me	3.22	2.38
	418506	AA084248	Hs.85339	G protein-coupled receptor 39	2.66	2.22
70	418526	BE019020	Hs.85838	solute carrier family 16 (monocarboxylic	2.04	2.21
	418538	BE244323	Hs.85951	exportin, tRNA (nuclear export receptor	1.33	37.00
	418543	NM_005329	Hs.85962	hyaluronan synthase 3	1.04	1.23
	418574	N28754		M-phase phosphoprotein 9	48.60	85.00
	418592	X99226	Hs.284153	Fanconi anemia, complementation group A	18.24	26.00
75	418641	BE243136	Hs.86947	a disintegrin and metalloproteinase doma	1.19	1.41
	418661	NM_001949	Hs.1189	E2F transcription factor 3	29.05	43.00
	418663	AK001100	Hs.41690	desmocollin 3	112.17	19.00
	418678	NM_001327	Hs.87225	cancer/testis antigen	1.18	1.10
	418686	Z36830	Hs.87268	annexin A8	1.54	1.98
80	418689	AI360883	Hs.274448	hypothetical protein FLJ11029	1.19	1.04
	418712	Z42183		gb:HSC0BF041 normalized infant brain cDN	1.00	12.00
	418727	AA227609	Hs.94834	ESTs	1.00	49.00
	418738	AW388633	Hs.6682	solute carrier family 7, (cationic amino	49.85	1.00
70	418819	AA228776	Hs.191721	ESTs	1.00	140.00
	418830	BE513731	Hs.88959	hypothetical protein MGC4816	20.97	23.00
85	418882	NM_004996	Hs.89433	ATP-binding cassette, sub-family C (CFTR	57.09	35.00

	418971	AA360392	Hs.87113	ESTs	1.00	12.00
	418973	AA233056	Hs.191518	ESTs	4.89	28.00
	419078	M93119	Hs.89584	insulinoma-associated 1	1.00	10.00
5	419079	AW014836	Hs.18344	ESTs	1.09	1.98
	419080	AW150835	Hs.18878	hypothetical protein FLJ21620	2.06	1.68
	419088	AJ538323	Hs.52620	integrin, beta 8	15.60	51.00
	419092	J05581	Hs.89603	mucin 1, transmembrane	1.11	1.83
	419121	AA374372	Hs.89626	parathyroid hormone-like hormone	1.00	1.00
10	419171	NM_002846	Hs.89655	protein tyrosine phosphatase, receptor t	1.10	1.14
	419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	1.00	1.00
	419216	AJ076718	Hs.164021	small inducible cytokine subfamily B (Cy	3.18	2.43
	419288	AA256106	Hs.87507	ESTs	1.00	34.00
	419335	AW960146	Hs.284137	hypothetical protein FLJ12888	1.00	8.00
	419354	M62839	Hs.1252	apolipoprotein H (beta-2-glycoprotein I)	22.63	54.00
15	419359	AL043202	Hs.90073	chromosome segregation 1 (yeast homolog)	2.50	1.98
	419423	D26488	Hs.90315	KIAA0007 protein	1.00	7.00
	419443	D62703		gb:HUM316G10B Clontech human aorta polyA	1.00	12.00
	419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	1.64	1.84
20	419474	AW968619	Hs.155849	ESTs	13.63	62.00
	419485	AA489023	Hs.99807	ESTs, Weakly similar to unnamed protein	4.27	2.26
	419488	AA316241	Hs.90691	nucleophosmin/nucleoplasm 3	3.66	3.63
	419502	AJ076704		fibrinogen, A alpha polypeptide	13.05	115.00
	419539	AF070590	Hs.90869	Homo sapiens clones 24622 and 24623 mRNA	74.60	117.00
25	419556	U29615	Hs.91093	chitinase 1 (chitinobiosidase)	1.47	4.98
	419569	AJ971651	Hs.91143	Jagged 1 (Alagille syndrome)	1.00	4.00
	419594	AA013051	Hs.91417	topoisomerase (DNA) II binding protein	94.30	94.00
	419703	AJ793257	Hs.128151	ESTs	15.26	50.00
	419721	NM_001650	Hs.288650	aquaporin 4	1.00	191.00
30	419729	AA586442	Hs.21411	gb:nc53a03.s1 NCL CGAP_SS1 Homo sapiens	1.00	59.00
	419741	NM_007019	Hs.93002	ubiquitin carrier protein E2-C	2.02	1.08
	419745	AF042001	Hs.93005	slug (chicken homolog), zinc finger prot	1.00	1.00
	419752	AA249573	Hs.152618	ESTs, Moderately similar to ZN91_HUMAN Z	29.87	77.00
	419839	U24577	Hs.93304	phospholipase A2, group VII (platelet-ac	50.99	214.00
35	419936	AJ792788		gb:cl91d05.y5 NCL CGAP_Kid5 Homo sapiens	1.00	1.00
	419937	AB040959	Hs.83836	DKFZP434N014 protein	1.64	2.47
	419983	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f	15.72	94.00
	420005	AW271106	Hs.133294	ESTs	3.15	1.43
	420047	AI478658	Hs.94631	brefeldin A-inhibited guanine nucleotide	12.45	39.00
40	420058	AK001423	Hs.94694	Homo sapiens cDNA FLJ10561 fis, clone NT	1.00	117.00
	420162	BE378432	Hs.95577	cyclin-dependent kinase 4	1.43	1.21
	420251	AW374968	Hs.348112	Human DNA sequence from clone RP5-1103G7	2.35	3.23
	420259	AF004884	Hs.96253	calcium channel, voltage-dependent, P/Q	0.77	1.15
	420281	AI623693	Hs.323494	ESTs	45.04	54.00
45	420309	AW043637	Hs.21766	ESTs, Weakly similar to ALU5_HUMAN ALU S	49.22	31.00
	420332	NM_001756	Hs.1305	serine (or cysteine) proteinase inhibitor	0.05	2.82
	420380	AA640891	Hs.102406	ESTs	0.99	2.74
	420462	AF050147	Hs.97932	chondromodulin I precursor	1.00	1.00
	420520	AK001978	Hs.98510	similar to rab11-binding protein	49.74	133.00
50	420552	AK000492	Hs.98806	hypothetical protein	94.65	88.00
	420560	AW207748	Hs.59115	ESTs	1.00	17.00
	420610	AI683183	Hs.99348	distal-less homeo box 5	1.00	13.00
	420689	H79979	Hs.886878	ESTs	50.09	95.00
	420721	AA927802	Hs.159471	ZAP3 protein	1.00	31.00
55	420759	T11832	Hs.127797	Homo sapiens cDNA FLJ11381 fis, clone HE	1.00	48.00
	420783	AI659838	Hs.99923	lectin, galactoside-binding, soluble, 7	3.04	1.25
	420900	AL045633	Hs.44269	ESTs	2.24	7.00
	420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	1.00	8.00
	421002	AF116030	Hs.100932	transcription factor 17	1.00	27.00
60	421027	AA761198	Hs.55254	ESTs	2.87	38.00
	421037	AI684808	Hs.197653	ESTs	1.00	46.00
	421041	N36914	Hs.14691	ESTs, Moderately similar to I38022 hypot	1.00	98.00
	421073	NM_004669	Hs.101448	metastasis associated 1	1.34	1.46
	421110	AJ250717	Hs.1355	cathepsin E	119.47	427.00
65	421133	AA401369	Hs.190721	ESTs	1.10	17.00
	421150	AI913562	Hs.189902	ESTs	1.45	1.63
	421155	H87879	Hs.102267	lysyl oxidase	1.00	15.00
	421307	BE539976	Hs.103305	Homo sapiens mRNA; cDNA DKFZp434B0425 (f	1.37	1.10
	421316	AA287203	Hs.324728	SMA5	1.00	21.00
70	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy	1.92	3.94
	421451	AA291377	Hs.50831	ESTs	5.89	14.00
	421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	1.46	1.76
	421506	BE302796	Hs.105097	thymidine kinase 1, soluble	1.56	1.08
	421508	NM_004833	Hs.105115	absent in melanoma 2	5.11	5.23
75	421515	Y11339	Hs.105352	GalNAc alpha-2, 6-sialyltransferase I, I	1.00	3.00
	421524	AA312082	Hs.105445	GDNF family receptor alpha 1	2.63	10.58
	421526	AL080121	Hs.105460	DKFZP564O0823 protein	1.46	1.88
	421552	AF026692	Hs.105700	secreted frizzled-related protein 4	30.21	50.32
	421574	AJ000152	Hs.105924	defensin, beta 2	1.67	1.74
80	421582	AI910275		trafallo factor 1 (breast cancer, estroge	1.23	1.00
	421633	AF121860	Hs.106260	sorting nexin 10	1.00	116.00
	421659	NM_014459	Hs.106511	protocadherin 17	0.05	6.33
	421677	H64092	Hs.38282	ESTs	1.31	1.42
	421753	BE314828	Hs.107911	ATP-binding cassette, sub-family B (MDR/	1.41	1.20
85	421773	W69233	Hs.112457	ESTs	1.12	1.14
	421777	BE562088	Hs.108196	HSPC037 protein	1.97	1.29

	421800	AA298151	Hs.222969	ESTs	1.03	1.30
	421817	AF146074	Hs.108560	ATP-binding cassette, sub-family C (CFTR	1.88	1.59
	421896	N62293	Hs.45107	ESTs	11.84	22.80
5	421928	AF013759	Hs.109643	polyadenylate binding protein-interactin	45.89	90.00
	421931	NM_000814	Hs.1440	gamma-aminobutyric acid (GABA) A recepto	1.13	1.49
	421948	L42583	Hs.334309	keratin 6A	51.83	20.25
	421975	AW961017	Hs.6459	hypothetical protein FLJ11856	1.17	1.15
	422026	U80736	Hs.110826	trinucleotide repeat containing 9	1.00	52.00
10	422094	AF129535	Hs.272027	F-box only protein 5	67.61	62.00
	422095	AI868872	Hs.282804	hypothetical protein FLJ22704	4.37	2.34
	422109	ST3265	Hs.1473	gastrin-releasing peptide	4.18	95.50
	422128	AW881145		gb:QV0-OT0033-010400-182-a07 OT0033 Homo	40.89	71.00
	422129	AU076635	Hs.1478	serine (or cysteine) proteinase inhibito	1.13	1.38
15	422134	AW179019	Hs.112110	mitochondrial ribosomal protein L42	41.59	95.00
	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	2.37	1.10
	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias	3.29	1.68
	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 6	4.93	5.73
	422282	AF019225	Hs.114309	apolipoprotein L	1.49	1.71
20	422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	25.99	10.91
	422310	AA316622	Hs.98370	cytochrome P450, subfamily IIS, polypept	1.54	1.41
	422311	AF073515	Hs.114948	cytokine receptor-like factor 1	1.15	1.78
	422330	D30783	Hs.115263	epiregulin	1.00	112.00
	422364	AF067800	Hs.115515	C-type (calcium dependent, carbohydrata-	9.39	60.00
25	422406	AF025441	Hs.116205	Opa-interacting protein 5	18.33	53.00
	422424	AI185431	Hs.296538	prostate differentiation factor	1.71	3.21
	422440	NM_004812	Hs.116724	aldo-keto reductase family 1, member B10	47.53	32.00
	422487	AJ010901	Hs.198267	mucin 4, tracheobronchial	73.68	35.54
	422511	AU076442	Hs.117938	collagen, type XVII, alpha 1	173.97	26.00
30	422515	AW500470	Hs.117950	multifunctional polypeptide similar to S	4.68	2.92
	422555	AJ870435	Hs.1569	LIM homeobox protein 2	1.00	1.00
	422737	M26939	Hs.119571	collagen, type III, alpha 1 (Ehlers-Danl	3.89	4.55
	422756	AA441787	Hs.119589	glycoprotein hormones, alpha polypeptide	1.05	1.46
	422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	3.88	1.53
35	422809	AK001379	Hs.121028	hypothetical protein FLJ10549	99.56	53.00
	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	1.69	3.17
	422938	NM_001809	Hs.1594	centromere protein A (17kD)	70.46	61.00
	422956	BE545072	Hs.122579	ECT2 protein (Epithelial cell transformi	77.74	3.00
	422960	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)	5.88	8.55
40	422963	AA401369	Hs.190721	ESTs	171.41	17.00
	422976	AU076657	Hs.1600	chaperonin containing TCP1, subunit 5 (e	2.12	1.62
	422981	AF026445	Hs.122752	TATA box binding protein (TBP)-associate	10.49	35.00
	422986	AA319777	Hs.221974	ESTs	12.40	32.47
	423034	AL119930		gb:DKFZp761A092_r1 761 (synonym: hamy2)	16.41	60.00
45	423049	X59373	Hs.188023	ESTs, Moderately similar to HXDA_HUMAN H	1.00	1.00
	423081	AF262992	Hs.123159	sperm associated antigen 4	1.82	2.96
	423184	NM_004428	Hs.1624	ephrin-A1	1.14	1.53
	423217	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys	2.14	1.69
	423248	AA380177	Hs.125845	ribulose-5-phosphate-3-epimerase	7.18	14.00
50	423309	BE006775	Hs.126782	sushi-repeat protein	21.90	64.00
	423361	AW170055	Hs.47628	ESTs	1.00	1.00
	423453	AW450737	Hs.128791	CGI-09 protein	55.52	66.00
	423511	AF036329	Hs.129715	gonadotropin-releasing hormone 2	0.88	1.17
	423516	AB007933	Hs.129729	ligand of neuronal nitric oxide synthase	1.76	5.40
55	423551	AA327598	Hs.233785	ESTs	3.54	4.33
	423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	1.00	50.00
	423575	C18863	Hs.163443	Homo sapiens cDNA FLJ11576 fis, clone HE	38.88	70.00
	423624	AI807408	Hs.166368	ESTs	1.00	67.00
	423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	76.02	1.00
60	423642	AW452650	Hs.157148	hypothetical protein MGC13204	19.14	58.00
	423662	AA642452	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro	3.61	13.57
	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	240.73	40.00
	423698	AA329796	Hs.1098	DKFZp434J1813 protein	1.00	59.00
	423725	AJ403108	Hs.132127	hypothetical protein LOC57822	4.20	1.00
65	423761	NM_006194	Hs.132576	paired box gene 9	1.00	1.00
	423787	AJ295745	Hs.236204	nuclear pore complex protein	7.18	6.64
	423816	AF151064		hypothetical protein	1.00	44.00
	423826	U20325	Hs.1707	cocaine- and amphetamine-regulated trans	1.00	1.00
	423849	AL157425	Hs.133315	Homo sapiens mRNA: cDNA DKFZp761J1324 (f	1.00	1.00
70	423887	AL080207	Hs.134585	DKFZP434G232 protein	1.00	1.00
	423934	U89995	Hs.159234	forkhead box E1 (thyroid transcription f	31.33	31.00
	423954	AW753184	Hs.288604	KIAA1632 protein	5.81	10.87
	423961	D13666	Hs.136348	osteoblast specific factor 2 (fasciclin	3.55	3.30
	424012	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	233.42	68.00
75	424016	AW163729	Hs.6140	hypothetical protein MGC15730	0.93	1.01
	424028	AF055084	Hs.153692	Homo sapiens cDNA FLJ14354 fis, clone Y7	21.30	52.00
	424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibito	1.00	1.00
	424086	AI351010	Hs.102267	lysyl oxidase	21.91	70.00
	424098	AF077374	Hs.139322	small proline-rich protein 3	137.82	54.00
80	424120	T80579	Hs.290270	ESTs	1.00	1.00
	424165	AW582904	Hs.142255	islet amyloid polypeptide	1.00	34.00
	424200	AA337221		gb:EST41944 Endometrial tumor Homo sapie	13.06	48.00
	424279	L29306	Hs.171814	tryptophan hydroxylase (tryptophan 5-mon	1.00	1.00
	424308	AW975531	Hs.154443	minichromosome maintenance deficient (S.	164.58	87.00
	424326	NM_014479	Hs.145296	disintegrin protease	53.72	302.00
85	424340	AA339036	Hs.7033	ESTs	0.88	1.15

	424351	BE622117	Hs.145567	hypothetical protein	0.93	1.03
	424364	AW383226	Hs.201189	ESTs, Weakly similar to G01763 atrophin-	7.02	3.24
	424381	AA285249	Hs.146329	protein kinase Chk2	95.55	92.00
5	424411	NM_005209	Hs.146549	crystallin, beta A2	1.63	3.25
	424420	BE614743	Hs.146588	prostaglandin E synthase	1.63	1.33
	424441	X14850	Hs.147097	H2A histone family, member X	1.82	1.29
	424502	AF242388	Hs.149585	lengsin	1.00	1.00
	424503	X06256	Hs.149509	integrin, alpha 5 (fibronectin receptor,	1.02	2.24
	424513	BE385864	Hs.149894	mitochondrial translational initiation f	1.00	17.00
10	424539	L02911	Hs.150402	Activin A receptor, type I (ACVR1) (ALK	32.46	108.00
	424568	AF005418	Hs.150595	cytochrome P450, subfamily XXVIA, polype	3.40	2.58
	424602	AK002055	Hs.151046	hypothetical protein FLJ11193	31.87	25.00
	424629	M90656	Hs.151393	glutamate-cysteine ligase, catalytic sub	3.58	2.37
	424645	NM_014682	Hs.151449	KIAA0535 gene product	1.00	1.00
15	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	2.12	2.23
	424717	AW992292	Hs.152213	wingless-type MMTV integration site fami	1.00	1.00
	424834	AK001432	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	56.19	12.00
	424840	D79987	Hs.153479	extra spindle poles, S. cerevisiae, homo	2.65	1.30
	424867	A024860	Hs.153591	Not56 (D. melanogaster)-like protein	1.23	1.05
20	424905	NM_002497	Hs.153704	NIMA (never in mitosis gene a)-related k	21.35	1.00
	424979	D87989	Hs.154073	UDP-galactose transporter related	1.36	1.35
	424999	AW953120		gb:EST365190 MAGE resequences, MAGB Homo	1.24	1.41
	425048	H05468	Hs.164502	ESTs	1.00	11.00
25	425057	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol	7.46	87.00
	425081	X74794	Hs.154443	minichromosome maintenance deficient (S	2.52	3.82
	425118	AU076611	Hs.154672	methylene tetrahydrofolate dehydrogenase	4.84	4.03
	425159	NM_004341	Hs.154868	carbamoyl-phosphate synthetase 2, aspart	3.62	2.73
	425202	AW962282	Hs.152049	ESTs, Weakly similar to I38022 hypothe	1.00	53.00
	425234	AW152225	Hs.165909	ESTs, Weakly similar to I38022 hypothe	100.77	44.00
30	425236	AW067800	Hs.155223	stanniocalcin 2	3.30	2.90
	425245	AJ751768	Hs.155314	KIAA0095 gene product	1.91	2.32
	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	1.41	1.49
	425266	J00077	Hs.155421	alpha-fetoprotein	1.00	68.00
	425274	BE281191	Hs.155462	minichromosome maintenance deficient (mi	1.97	1.63
35	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	141.49	123.00
	425349	AA425234	Hs.79886	ribose 5-phosphate isomerase A (ribose 5	1.00	84.00
	425371	D49441	Hs.155981	mesothelin	0.87	1.59
	425397	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	14.90	5.76
40	425420	BE536911	Hs.234545	hypothetical protein NUF2R	1.00	1.00
	425424	NM_004954	Hs.157199	ELKL motif kinase	10.58	9.74
	425483	AF231022	Hs.158159	FAT tumor suppressor (Drosophila) homolo	1.74	1.40
	425566	AW162943	Hs.250618	UL16 binding protein 2	1.49	1.14
	425580	L11144	Hs.1907	galanin	53.29	233.00
45	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	33.45	1.00
	425692	D90041	Hs.155956	N-acetyltransferase 1 (arylamine N-acety	1.00	55.00
	425695	NM_005401	Hs.159238	protein tyrosine phosphatase, non-recept	1.00	10.00
	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg	1.00	41.00
	425776	U25128	Hs.159499	parathyroid hormone receptor 2	1.00	48.00
50	425810	AJ923627	Hs.31903	ESTs	27.39	98.00
	425811	AL039104	Hs.155557	karyopherin alpha 2 (RAG cohort 1, impor	1.99	1.58
	425849	AI077288	Hs.296323	serum/glucocorticoid regulated kinase	71.16	3.42
	425852	AK001504	Hs.158661	death receptor 6, TNF superfamily member	1.35	1.34
	426067	AA401369	Hs.190721	ESTs	1.01	17.00
55	426088	AF038007	Hs.166195	ATPase, Class I, type 8B, member 1	26.26	47.00
	426215	AW067800	Hs.155223	stanniocalcin 2	1.91	2.90
	426227	U67058	Hs.154299	Human proteinase activated receptor-2 mR	22.40	25.00
	426269	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f	1.00	1.00
	426283	NM_003937	Hs.169139	kynureninase (L-kynurenine hydrolase)	91.39	229.00
60	426329	AL389951	Hs.271623	nucleoporin 50kD	4.34	4.08
	426427	M85699	Hs.169840	TTK protein kinase	7.02	1.00
	426432	AF001601	Hs.169887	paraoxonase 2	1.16	1.68
	426440	BE382756	Hs.169902	solute carrier family 2 (facilitated glu	2.59	1.71
	426459	AF151812	Hs.169992	hypothetical 43.2 Kd protein	1.56	1.66
	426471	M22440	Hs.170009	transforming growth factor, alpha	20.60	26.00
65	426496	D31765	Hs.170114	KIAA0061 protein	9.81	22.00
	426501	AA401369	Hs.190721	ESTs	19.23	17.00
	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic	103.74	41.00
	426536	AJ949749	Hs.44441	ESTs	4.65	23.00
	426572	AB037783	Hs.170623	hypothetical protein FLJ11183	1.00	43.00
70	426682	AV660038	Hs.2056	UDP glycosyltransferase 1 family, polype	160.08	8.00
	426691	NM_006201	Hs.171834	PCTAIRE protein kinase 1	1.51	1.35
	426746	J03626	Hs.2057	uridine monophosphate synthetase (orotid	2.13	1.68
	426752	X69490	Hs.172004	tin	0.02	5.14
	426784	U03749	Hs.172216	chromogranin A (parathyroid secretory pr	1.72	1.71
75	426807	AA385315	Hs.156682	ESTs	1.30	1.64
	426812	AF105365	Hs.172613	solute carrier family 12 (potassium/chlo	1.47	1.53
	426814	AF036943	Hs.172619	myelin transcription factor 1-like	1.00	1.00
	426831	BE296216	Hs.172673	S-adenosylhomocysteine hydrolase	1.51	1.25
	426897	AA401369	Hs.190721	ESTs	141.56	17.00
80	426925	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	32.61	38.00
	426935	NM_000088	Hs.172928	collagen, type I, alpha 1	2.65	3.16
	426964	AA393739	Hs.287416	Homo sapiens cDNA FLJ11439 fis, clone HE	1.97	3.49
	426966	AI493134		sclerostin	1.00	1.00
	426991	AK001536		Homo sapiens cDNA FLJ10574 fis, clone NT	3.39	2.28
85	427099	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	4.24	17.00

	427239	BE270447	Hs.174070	ubiquitin carrier protein	1.58	1.05
	427260	AA663848		gb:ae70b06.s1 Stralagene schizo brain S1	1.34	1.60
	427281	AA906147	Hs.102869	ESTs	1.00	66.00
5	427335	AA448542	Hs.251877	G antigen 7B	51.83	4.00
	427354	T57896	Hs.191095	ESTs	1.17	1.95
	427356	AW023482	Hs.97849	ESTs	7.31	41.00
	427356	AA401533	Hs.19440	ESTs	1.00	57.00
	427383	NM_005411	Hs.177582	surfactant, pulmonary-associated protein	0.42	1.32
10	427427	AF077345	Hs.177936	lectin, superfamily member 1 (cartilage-	1.00	20.00
	427441	AA412605	Hs.343879	SPANX family, member C	1.00	1.00
	427445	X80818	Hs.178078	glutamate receptor, metabotropic 4	0.97	1.03
	427505	AA361562	Hs.178761	26S proteasome-associated pad1 homolog	4.60	4.04
	427510	Z47542	Hs.179312	small nuclear RNA activating complex, po	22.00	45.00
15	427528	AU077143	Hs.179565	minichromosome maintenance deficient (S.	97.45	92.00
	427546	AA189763	Hs.36793	hypothetical protein FLJ23188	1.50	3.24
	427562	R56424	Hs.26534	ESTs	6.81	40.00
	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	69.91	62.00
	427660	AJ741320	Hs.114121	Homo sapiens cDNA: FLJ23228 fis, clone C	2.70	49.00
20	427666	AJ791495	Hs.180142	calmodulin-like skin protein	1.37	1.88
	427668	AA298760	Hs.180191	hypothetical protein FLJ14904	29.55	67.00
	427677	NM_007045	Hs.180295	FGFR1 oncogene partner	3.52	2.63
	427701	AA411101	Hs.243886	nuclear autoantigenic sperm protein (his	7.41	34.00
	427711	M31659	Hs.180408	solute carrier family 25 (mitochondrial	15.84	70.00
	427719	AJ393122	Hs.134726	ESTs	7.03	4.52
25	427722	AK000123	Hs.180479	hypothetical protein FLJ20116	2.92	1.74
	427747	AW411425	Hs.180555	serine/threonine kinase 12	1.76	1.26
	427912	AL022310	Hs.181097	tumor necrosis factor (ligand) superfam	9.63	59.00
	427961	AW293165	Hs.143134	ESTs	41.97	118.00
30	428004	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub	23.82	1.00
	428023	AL038843		Homo sapiens cDNA: FLJ23602 fis, clone L	1.40	1.33
	428046	AW812795	Hs.337534	ESTs, Moderately similar to I36022 hypot	96.28	167.00
	428093	AW594506	Hs.104830	ESTs	1.25	1.29
	428098	AU077258	Hs.182429	protein disulfide isomerase-related prot	1.86	1.60
35	428129	AJ244311	Hs.26912	ESTs	1.00	42.00
	428169	AI928984	Hs.182793	golgi phosphoprotein 2	2.76	2.11
	428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN G ANT	1.00	1.00
	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy	85.59	181.00
	428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	8.57	21.64
40	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	7.77	15.90
	428434	AI909935	Hs.65551	Homo sapiens, Similar to DNA segment, Ch	0.58	1.43
	428450	NM_014791	Hs.184339	KIAA0175 gene product	237.53	204.00
	428471	X57348	Hs.184510	stratificin	6.00	4.60
	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to	56.54	16.00
45	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino	3.53	2.15
	428505	AL035461	Hs.2281	chromogranin B (secretogranin 1)	1.00	1.00
	428532	AF157326	Hs.184786	TBP-interacting protein	1.00	58.00
	428545	AA431400	Hs.98729	ESTs, Weakly similar to 2017205A dihydro	1.00	16.00
	428664	AK001666	Hs.189095	similar to SALL1 (sal (Drosophila)-like	1.00	1.00
50	428698	AA852773	Hs.334838	KIAA1866 protein	187.37	255.00
	428728	NM_016625	Hs.191381	hypothetical protein	47.24	80.00
	428748	AW593206	Hs.98785	Ksp37 protein	1.00	87.00
	428758	AA433988	Hs.98502	hypothetical protein FLJ14303	1.06	1.13
	428771	AB028992	Hs.193143	KIAA1069 protein	1.98	92.00
55	428801	AW277121	Hs.254881	ESTs	1.67	6.15
	428810	AF068236	Hs.193788	nitric oxide synthase 2A (inducible, hep	1.03	1.27
	428839	AI767756	Hs.82302	Homo sapiens cDNA FLJ14814 fis, clone NT	124.17	43.00
	428845	AL157579	Hs.153610	KIAA0751 gene product	1.00	1.00
	428959	AF100779	Hs.194680	WNT1 inducible signaling pathway protein	15.16	27.00
60	428969	AF120274	Hs.194689	artemin	1.36	1.24
	429038	AL023513	Hs.194766	seizure related gene 6 (mouse)-like	0.97	3.31
	429055	AI753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	6.82	16.47
	429164	AI688663	Hs.116586	ESTs	19.08	67.00
	429170	NM_001394	Hs.2359	dual specificity phosphatase 4	16.18	105.00
65	429183	AB014604	Hs.197955	KIAA0704 protein	79.72	104.00
	429201	X03178	Hs.198246	group-specific component (vitamin D bind	1.00	1.00
	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	1.33	1.09
	429220	AW207206		ESTs	1.00	7.00
	429228	AI553633	Hs.326447	ESTs	39.47	29.25
70	429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	2.01	1.18
	429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	1.07	1.00
	429276	AF056085	Hs.198612	G protein-coupled receptor 51	3.70	142.00
	429359	W00482	Hs.2399	matrix metalloproteinase 14 (membrane-in	1.30	1.94
	429412	NM_006235	Hs.2407	POU domain, class 2, associating factor	94.09	86.00
75	429413	NM_014058	Hs.201877	DESC1 protein	41.91	10.00
	429486	AF155827	Hs.203963	hypothetical protein FLJ10339	12.19	1.00
	429504	X99133	Hs.204238	lipocalin 2 (oncogene 24p3)	1.61	1.08
	429538	BE182592	Hs.11261	small proline-rich protein 2A	4.43	2.90
	429547	AA401369	Hs.190721	ESTs	1.06	17.00
80	429551	AW450624	Hs.220931	ESTs	2.89	65.00
	429563	BE619413	Hs.2437	eukaryotic translation initiation factor	1.49	1.37
	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	61.86	100.00
	429610	AB024937	Hs.211092	LUNX protein; PLLNC (palate lung and nas	1.59	1.69
	429612	AF062649	Hs.252587	pituitary tumor-transforming 1	2.78	1.74
85	429616	AI982722	Hs.120846	ESTs	1.00	1.00
	429656	X05608	Hs.211584	neurofilament, light polypeptide (68kD)	1.00	4.00

	429663	M58874	Hs.211587	phospholipase A2, group IVA (cytosolic,	69.95	104.00
	429736	AF125304	Hs.212680	tumor necrosis factor receptor superfamily	1.25	1.21
	429782	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain	1.00	7.00
5	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	11.80	1.00
	429918	AW873986	Hs.119383	ESTs	1.00	78.00
	429978	AA249027		ribosomal protein S6	1.98	3.09
	429986	AF092047	Hs.227277	sine oculis homeobox (Drosophila) homolo	1.00	48.00
	430044	AA464510	Hs.152812	ESTs	69.27	59.00
10	430114	AA847744	Hs.99640	ESTs	1.00	1.00
	430134	BE380149	Hs.105223	ESTs, Weakly similar to T33188 hypothe	1.00	51.00
	430147	R60704	Hs.234434	hair/enhancer-of-split related with YRP	1.10	2.22
	430287	AW182459	Hs.125759	ESTs, Weakly similar to LEU5_HUMAN LEUKE	1.00	127.00
	430294	AJ538226	Hs.32976	guanine nucleotide binding protein 4	3.80	1.47
15	430300	U60805	Hs.238648	oncostatin M receptor	1.00	35.00
	430315	NM_004293	Hs.239147	guanine deaminase	92.31	28.00
	430337	M35707	Hs.239600	calmodulin-like 3	1.18	1.08
	430378	Z29572	Hs.2555	tumor necrosis factor receptor superfamily	5.28	66.00
	430388	AA356923	Hs.240770	nuclear cap binding protein subunit 2, 2	16.76	38.00
	430393	BE185030	Hs.241305	estrogen-responsive B box protein	1.63	1.50
20	430439	AL133561		DKFZP434B061 protein	1.00	1.00
	430451	AA836472	Hs.297939	cathepsin B	1.64	2.12
	430454	AW469011	Hs.105635	ESTs	63.35	44.00
	430466	AF052573	Hs.241517	polymerase (DNA directed), theta	2.47	1.91
25	430481	AA479678	Hs.203269	ESTs, Moderately similar to ALU8_HUMAN A	1.00	31.00
	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	12.28	41.00
	430508	AJ015435	Hs.104637	ESTs	4.75	7.27
	430533	AA480895	Hs.57749	ESTs, Weakly similar to T17288 hypothe	1.00	1.00
	430563	AF145074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	1.00	1.59
30	430577	Z26317	Hs.94560	desmoglein 2	1.72	1.30
	430578	AA401369	Hs.190721	ESTs	0.90	17.00
	430686	NM_001942	Hs.2633	desmoglein 1	1.00	1.00
	430788	AJ742925	Hs.7179	ESTs, Weakly similar to 2004399A chromos	1.62	1.84
	430890	X54232	Hs.2699	glypican 1	1.58	1.40
35	430935	AW072916		zinc finger protein 131 (clone pHZ-10)	90.28	132.00
	430985	AA490232	Hs.27323	ESTs, Weakly similar to I78885 serine/th	0.94	1.28
	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	60.25	28.00
	431089	BE041395		ESTs, Weakly similar to unknown protein	23.32	941.00
40	431092	AJ332764	Hs.125757	ESTs	13.46	63.00
	431124	AF284221	Hs.59506	doublesex and mab-3 related transcrip	49.43	62.00
	431164	AA493650	Hs.94367	Homo sapiens cDNA: FLJ23494 fis, clone L	0.44	2.20
	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	182.26	101.00
	431221	AW207837	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye	4.15	13.97
45	431277	AA501806	Hs.345824	ESTs	1.00	86.00
	431322	AW970622		gb:EST382704 MAGE resequences, MAGK Homo	40.55	200.00
	431342	AW971018	Hs.21659	ESTs	1.00	53.00
	431384	BE158000	Hs.285026	gb:MR2-HT0377-150203-202-e03 HT0377 Homo	0.94	1.14
	431462	AW53672	Hs.256311	granin-like neuroendocrine peptide precu	1.30	1.25
	431494	AA991355	Hs.298312	hypothetical protein DKFZp434A1315	3.90	26.00
50	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha	1.41	1.87
	431548	AJ834273	Hs.9711	novel protein	5.66	15.00
	431630	NM_002204	Hs.265829	integrin, alpha 3 (antigen CD49C, alpha	0.99	1.44
	431745	AW972448	Hs.163425	ESTs	0.99	3.51
	431770	BE221880	Hs.268555	5'-3' exonuclease 2	67.12	91.00
55	431830	Y16645	Hs.271387	small inducible cytokine subfamily A (Cy	3.36	4.71
	431846	BE019924	Hs.271580	uroplakin 1B	4.49	2.51
	431890	X17033	Hs.271986	integrin, alpha 2 (CD49B, alpha 2 subuni	2.20	3.32
	431934	AB031481	Hs.272214	STG protein	1.01	1.04
60	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	51.17	46.35
	432006	AL137382	Hs.272320	Homo sapiens mRNA; cDNA DKFZp434L1226 (f	0.94	1.65
	432023	R43020	Hs.236223	EST	0.94	47.00
	432201	AJ538613	Hs.298241	Transmembrane protease, serine 3	1.10	2.24
	432210	AJ567421	Hs.273330	Homo sapiens, clone IMAGE:3544662, mRNA,	1.42	1.45
65	432226	AW182766	Hs.273558	phosphate cytidylyltransferase 1, cholin	1.00	1.00
	432239	X81334	Hs.2936	matrix metalloproteinase 13 (collagenase	18.67	1.00
	432265	BE382679	Hs.285753	SCG10-like-protein	1.09	1.21
	432281	AK001239	Hs.274263	hypothetical protein FLJ10377	40.98	58.00
	432365	AK001106	Hs.274419	hypothetical protein FLJ10244	1.00	214.00
	432374	W68815	Hs.301885	Homo sapiens cDNA FLJ11346 fis, clone PL	157.34	37.00
70	432375	BE536069	Hs.2962	S100 calcium-binding protein P	1.65	1.06
	432407	AA221036		gb:zr03f12.r1 Stratagene NT2 neuronal pr	73.71	75.00
	432441	AW292425	Hs.163484	ESTs	56.35	72.00
	432489	AJ804855	Hs.207530	ESTs	1.00	24.00
	432543	AA552690	Hs.152423	Homo sapiens cDNA: FLJ21274 fis, clone C	137.72	98.00
75	432552	AJ537170	Hs.173725	ESTs, Weakly similar to ALU8_HUMAN ALU S	1.00	31.00
	432583	AW023624	Hs.162282	potassium channel TASK-4; potassium chan	0.27	35.18
	432606	NM_002104	Hs.3066	granzyme K (serine protease, granzyme 3;	2.87	6.22
	432625	AJ243596	Hs.94830	ESTs, Moderately similar to T03094 A-kin	26.63	56.00
	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	1.92	5.29
80	432677	NM_004482	Hs.276611	UDP-N-acetyl-alpha-D-galactosamine:polyp	1.00	48.00
	432715	AA247152	Hs.200483	ESTs, Weakly similar to KIAA1074 protein	45.13	31.00
	432753	NM_014075	Hs.336938	Homo sapiens PRO0593 mRNA, complete cds	1.00	68.00
	432788	AA521091	Hs.178499	Homo sapiens cDNA: FLJ23117 fis, clone L	2.69	3.67
	432842	AW674093	Hs.334822	hypothetical protein MGC4485	1.22	1.34
85	432867	AW016936	Hs.233364	ESTs	1.00	1.00
	432917	NM_014125	Hs.241517	PRO0327 protein	10.25	6.62

	432920	U37689	Hs.3128	polymerase (RNA) II (DNA directed) polyp	1.44	1.30
	433001	AF217513	Hs.279905	clone HQ0310 PRO0310p1	154.79	85.64
	433023	AW864793	Hs.87409	thrombospondin 1	20.95	100.00
5	433042	AW193534	Hs.281895	Homo sapiens cDNA FLJ11660 fis, clone HE	1.00	10.00
	433091	Y12642	Hs.3185	lymphocyte antigen 6 complex, locus D	1.20	1.09
	433159	AB035898	Hs.150587	kinesin-like protein 2	13.82	39.00
	433183	AF231338	Hs.222024	transcription factor BMAL2	1.00	69.00
	433258	AA622788	Hs.203613	ESTs, Weakly similar to ALUB_HUMAN III	1.00	1.25
	433409	AI278802	Hs.25661	ESTs	44.81	117.00
10	433437	U20536	Hs.3280	caspase 6, apoptosis-related cysteine pr	70.39	105.00
	433485	AI93076	Hs.201967	aldo-keto reductase family 1, member C2	11.55	2.00
	433537	AI733692	Hs.112488	ESTs	8.65	55.00
	433547	W04978	Hs.303023	beta tubulin 1, class VI	25.16	83.00
	433556	W56321	Hs.111460	calcium/calmodulin-dependent protein kin	1.00	19.00
15	433647	AA603367	Hs.222294	ESTs	20.30	49.00
	433658	L03678	Hs.156110	immunoglobulin kappa constant	5.92	10.03
	433800	AI094221	Hs.135150	lung type-I cell membrane-associated gly	2.29	2.22
	433819	AW511097	Hs.112765	ESTs	3.71	8.00
	433862	D86960	Hs.3610	KIAA0205 gene product	62.08	104.00
20	433980	AA137152	Hs.286049	phosphoserine aminotransferase	108.91	47.00
	434088	AF116677	Hs.249270	hypothetical protein PRO1956	1.00	1.00
	434094	AA305599	Hs.238205	hypothetical protein PRO2013	121.27	87.00
	434105	AW952124	Hs.13094	presenilins associated rhomboid-like pro	1.22	1.23
	434217	AW014795	Hs.23349	ESTs	14.11	57.00
25	434340	AI193043	Hs.128685	ESTs, Weakly similar to T17226 hypotheti	2.10	2.56
	434360	AA401369	Hs.190721	ESTs	40.98	17.00
	434414	AI798376		gb:tr34b07.x1 NCI_CGAP_Ov23 Homo sapiens	1.48	1.56
	434424	AI811202	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	1.00	64.00
30	434467	BE552368	Hs.231853	Homo sapiens cDNA FLJ13445 fis, clone PL	54.91	85.00
	434551	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	2.46	2.00
	434627	AI221894	Hs.39311	ESTs	1.00	1.00
	434699	AA643687	Hs.149425	Homo sapiens cDNA FLJ11980 fis, clone HE	1.00	23.00
	434769	AA648884	Hs.134278	Homo sapiens cDNA FLJ12676 fis, clone NT	7.08	56.00
35	434792	AA549253	Hs.132458	ESTs	8.52	44.00
	434808	AF155108	Hs.256150	Homo sapiens, Similar to RIKEN cDNA 2810	11.33	1.00
	434828	D90070	Hs.96	phorbol-12-myristate-13-acetate-induced	1.00	1.00
	434876	AF160477	Hs.61460	lg superfamily receptor LNIR	1.25	1.29
	434891	AA814309	Hs.123583	ESTs	1.00	6.00
40	434928	AW015595	Hs.4267	Homo sapiens clones 24714 and 24715 mRNA	1.00	1.00
	435013	H91923	Hs.110024	Target CAT	1.26	1.10
	435066	BE261750	Hs.4747	dyskeratosis congenita 1, dyskerin	1.69	1.37
	435087	AW975241	Hs.23567	ESTs	1.00	1.00
	435099	AC004770	Hs.4756	flap structure-specific endonuclease 1	2.90	1.93
45	435159	AA668879	Hs.116649	ESTs	1.00	1.00
	435205	X54136	Hs.181125	immunoglobulin lambda locus	1.02	1.46
	435232	NM_001262	Hs.4854	cyclin-dependent kinase inhibitor 2C (p1	2.04	2.70
	435304	H10709	Hs.269524	ESTs	27.58	139.00
	435313	AI769400	Hs.189729	ESTs	1.00	14.00
50	435505	AF200492	Hs.211238	interleukin-1 homolog 1	1.00	38.00
	435509	AI458679	Hs.181915	ESTs	1.00	1.00
	435525	AI831297	Hs.123310	ESTs	1.00	56.00
	435532	AW291488	Hs.117305	Homo sapiens, clone IMAGE:3682908, mRNA	1.00	2.00
	435550	AI224456	Hs.324507	H.sapiens polyA site DNA	3.42	3.92
55	435602	AF217515	Hs.283532	uncharacterized bone marrow protein BM03	3.95	1.80
	435766	R11673	Hs.186498	ESTs	1.00	28.00
	435793	AB037734	Hs.4993	KIAA1313 protein	23.68	42.00
	436069	AI056879	Hs.263209	ESTs	1.00	58.00
	436170	AW450381	Hs.14529	ESTs	1.00	18.00
60	436211	AK001581	Hs.334828	hypothetical protein FLJ10719; KIAA1794	5.84	22.00
	436213	AA325512	Hs.71472	hypothetical protein FLJ10774; KIAA1709	1.42	1.27
	436217	T53925	Hs.107	fibrinogen-like 1	57.97	31.00
	436238	AK002163	Hs.301724	hypothetical protein FLJ11301	2.51	1.71
	436251	BE515065	Hs.296585	nucleolar protein (KKE/D repeat)	2.33	1.64
65	436291	BE568452	Hs.344037	protein regulator of cytokinesis 1	108.99	52.00
	436302	AL355841	Hs.99330	hypothetical protein FLJ23588	0.75	2.81
	436396	AW992292	Hs.152213	wingless-type MMTV Integration site fami	60.01	1.00
	436414	BE264633	Hs.143638	WD repeat domain 4	2.50	2.19
	436419	AI948626	Hs.171356	ESTs	0.95	1.33
	436443	AW138211	Hs.128748	ESTs	1.12	9.26
70	436474	AJ270693	Hs.199887	ESTs	1.00	1.00
	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	3.28	1.56
	436486	AA742221	Hs.120633	ESTs	1.00	19.00
	436511	AA721252	Hs.291502	ESTs	16.76	14.00
75	436553	X57809	Hs.181125	immunoglobulin lambda locus	1.08	1.74
	436557	W15573	Hs.5027	ESTs, Weakly similar to A47582 B-cell gr	19.20	9.75
	436608	AA628980		down syndrome critical region protein DS	33.92	25.00
	436667	AW025183	Hs.127680	ESTs	0.89	1.19
	436771	AW975687	Hs.292979	ESTs	1.00	10.00
	436839	AA401369	Hs.190721	ESTs	1.00	17.00
80	436887	AW953157	Hs.193235	hypothetical protein DKFZp547D155	1.05	1.15
	436944	AW268614	Hs.5840	ESTs	1.00	1.00
	436961	AW375974	Hs.156704	ESTs	25.13	25.00
	436972	AA284679	Hs.25640	claudin 3	1.59	1.46
	437016	AI076916	Hs.5398	guanine monophosphate synthetase	2.35	1.78
85	437044	AL035864	Hs.69517	cDNA for differentially expressed CO16 g	1.34	1.13

	437181	AI306615	Hs.125343	ESTs, Weakly similar to KIAA0758 protein	1.00	17.00
	437204	AL110216	Hs.22826	ESTs, Weakly similar to I55214 salivary	40.55	82.00
	437205	AL110232	Hs.279243	Homo sapiens mRNA; cDNA DKFZp564D2071 (f	1.00	112.00
	437259	AI377755	Hs.120695	ESTs	1.00	205.00
5	437270	R18087	Hs.323769	cisplatin resistance related protein CRR	1.56	1.54
	437271	AL137445	Hs.28846	Homo sapiens mRNA; cDNA DKFZp566O134 (fr	113.25	125.00
	437370	AL359567	Hs.161952	Homo sapiens mRNA; cDNA DKFZp547D023 (fr	1.82	4.57
	437390	AI125859	Hs.112607	ESTs	1.35	1.75
10	437412	BE069288	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (fr	3.58	3.20
	437435	AI306152	Hs.27027	hypothetical protein DKFZp762H1311	3.03	1.08
	437444	HA5008	Hs.31518	ESTs	1.00	39.00
	437558	AI954795	Hs.156135	ESTs	1.00	19.00
	437623	D63880	Hs.5719	chromosome condensation-related SMC-asso	1.95	1.57
15	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti	1.00	3.00
	437814	AI088192	Hs.135474	ESTs, Weakly similar to DDX9_HUMAN ATP-D	1.00	45.00
	437840	AA884835	Hs.292014	ESTs	1.07	1.78
	437852	BE001836	Hs.256897	ESTs, Weakly similar to DJ365O12.1 [H.s.a	1.68	3.26
	437879	BE282082	Hs.5894	hypothetical protein FLJ10305	1.87	2.52
20	437915	AI637993	Hs.202312	Homo sapiens clone N11 Ntera2D1 leratoca	74.05	35.00
	437916	BE565249	Hs.20999	hypothetical protein FLJ23142	23.15	89.00
	437937	AI917222	Hs.121655	ESTs	1.00	1.00
	437942	AI888256	Hs.307526	ESTs	12.28	31.00
	438091	AW373062		nuclear receptor subfamily 1, group I, m	1.53	10.85
25	438113	AI467908	Hs.8882	ESTs	1.80	2.39
	438119	AW963217	Hs.203961	ESTs, Moderately similar to AF116721 89	22.67	36.90
	438274	AI918906	Hs.55080	ESTs	1.00	1.00
	438378	AW970529	Hs.86434	hypothetical protein FLJ21816	38.92	38.00
	438403	AA806607	Hs.292206	ESTs	1.00	1.00
30	438494	AA908678	Hs.130183	ESTs	2.05	80.00
	438546	AW297204	Hs.125811	ESTs	1.00	131.00
	438552	AJ245820	Hs.6314	type I transmembrane receptor (seizure-r	1.43	1.45
	438702	AI879064	Hs.54618	ESTs	1.00	34.00
	438724	AW612553	Hs.114670	Human DNA sequence from clone RP11-16L21	1.33	1.10
35	438746	AI885815	Hs.184727	Human melanoma-associated antigen p97 (m	2.42	1.59
	438779	NM_003787	Hs.6414	nucleolar protein 4	1.00	18.00
	438821	AA826425	Hs.192375	ESTs	2.03	2.57
	438885	AI886558	Hs.184987	ESTs	6.42	88.00
	438898	AA401369	Hs.190721	ESTs	22.41	17.00
40	438915	AA280174	Hs.285881	Williams-Beuren syndrome chromosome regi	1.00	1.00
	438956	W00847	Hs.135056	Human DNA sequence from clone RP5-850E9	2.20	1.88
	439000	AW979121		gb:EST391231 MAGE resequences, MAGP Homo	2.78	4.81
	439023	AA745978	Hs.28273	ESTs	1.17	1.31
	439024	R95696	Hs.35598	ESTs	1.00	28.00
45	439128	AI949371	Hs.153089	ESTs	1.00	67.00
	439146	AW138909	Hs.156110	immunoglobulin kappa constant	1.38	1.41
	439223	AW238299	Hs.250618	UL16 binding protein 2	1.93	1.64
	439285	AL133916		hypothetical protein FLJ20093	46.23	139.00
50	439318	AW837046	Hs.6527	G protein-coupled receptor 56	2.00	2.20
	439343	AF086161	Hs.114611	hypothetical protein FLJ11808	6.10	7.37
	439394	AA401369	Hs.190721	ESTs	3.39	17.00
	439410	AA632012	Hs.188746	ESTs	1.83	3.07
	439451	AF086270	Hs.278554	heterochromatin-like protein 1	23.28	52.00
	439452	AA918317	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro	18.76	122.00
55	439453	BE264974	Hs.6566	thyroid hormone receptor interactor 13	2.78	1.58
	439477	W69813	Hs.58042	ESTs, Moderately similar to GFR3_HUMAN G	1.22	1.44
	439492	AF086310	Hs.103159	ESTs	7.46	39.00
	439523	W72348	Hs.185029	ESTs	1.00	1.19
	439592	AF086413	Hs.58399	ESTs	1.00	1.00
60	439606	W79123	Hs.58561	G protein-coupled receptor 87	33.61	1.00
	439670	AF088076	Hs.59507	ESTs, Weakly similar to AC004858 3 U1 sm	1.00	1.00
	439702	AW085525	Hs.134182	ESTs	4.30	10.00
	439705	AW872527	Hs.59761	ESTs, Weakly similar to DAPI_HUMAN DEATH	86.55	11.00
	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	2.36	1.88
65	439750	AL359053	Hs.57564	Homo sapiens mRNA full length insert cDN	2.02	6.08
	439759	AL359055	Hs.67709	Homo sapiens mRNA full length insert cDN	1.00	21.00
	439780	AL109688		gb:Homo sapiens mRNA full length insert	7.27	25.00
	439840	AW449211	Hs.105445	GDNF family receptor alpha 1	1.00	1.00
	439926	AW014875	Hs.137007	ESTs	32.58	71.00
70	439963	AW247529	Hs.6793	platelet-activating factor acetylhydrola	21.28	9.55
	439979	AW600291	Hs.6823	hypothetical protein FLJ10430	68.83	61.00
	440006	AK000517	Hs.6844	hypothetical protein FLJ20510	1.83	4.02
	440028	AW473675	Hs.125843	ESTs, Weakly similar to T17227 hypotheti	1.42	2.54
	440106	AA864968	Hs.127699	KIAA1603 protein	1.00	54.00
	440138	AB033023	Hs.318127	hypothetical protein FLJ10201	24.18	52.00
75	440273	AI805392	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	3.21	4.72
	440289	AW450991	Hs.192071	ESTs	38.63	113.00
	440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma	62.88	147.00
	440492	R39127	Hs.21433	hypothetical protein DKFZp547J036	2.35	3.62
	440527	AV657117	Hs.184164	ESTs, Moderately similar to S65657 alpha	10.84	57.00
80	440659	AF134160	Hs.7327	claudin 1	3.18	2.37
	440704	M69241	Hs.162	insulin-like growth factor binding prote	2.89	2.09
	440943	AW082293	Hs.146161	hypothetical protein MGC2408	2.02	1.41
	440984	AI160011	Hs.272068	ESTs	1.29	1.14
	441020	AA401369	Hs.190721	ESTs	142.99	17.00
85	441031	AI110684	Hs.7645	fibrinogen, B beta polypeptide	1.41	99.00

	441128	AA570256	ESTs, Weakly similar to T23273 hypotheti	4.13	3.50
	441290	W27501	Hs.89605 cholinergic receptor, nicotinic, alpha p	1.00	1.00
	441352	BE614410	Hs.23044 RAD51 (S. cerevisiae) homolog (E coli Re	130.23	43.00
	441377	BE218239	Hs.202656 ESTs	22.03	1.00
5	441390	AI692560	Hs.131175 ESTs	3.65	7.70
	441497	R51064	Hs.23172 ESTs	1.00	1.00
	441525	AW241867	Hs.127728 ESTs	1.53	1.42
	441553	AA281219	Hs.121295 ESTs	1.89	1.57
	441607	NM_005010	Hs.7912 neuronal cell adhesion molecule	1.47	2.11
10	441633	AW958544	Hs.112242 normal mucosa of esophagus specific 1	216.22	363.00
	441636	AA081846	Hs.7921 Homo sapiens mRNA; cDNA DKFZp566E183 (fr	2.31	2.05
	441737	X79449	Hs.7957 adenosine deaminase, RNA-specific	1.30	1.49
	441790	AA401369	Hs.190721 ESTs	44.15	17.00
	441801	AW242799	Hs.86366 ESTs	1.00	1.00
15	441919	AI553802	Hs.128121 ESTs	1.00	122.00
	441937	R41782	Hs.22279 ESTs	0.86	1.37
	441954	AI744935	Hs.8047 Fanconi anemia, complementation group G	1.48	1.39
	442025	AW887434	Hs.11810 CDA11 protein	1.00	46.00
20	442029	AW956698	Hs.14456 neural precursor cell expressed, develop	9.92	45.00
	442072	AI740832	Hs.12311 Homo sapiens clone 23570 mRNA sequence	25.05	77.00
	442108	AW452649	Hs.166314 ESTs	3.61	3.14
	442117	AW664964	Hs.128899 ESTs	3.00	5.49
	442137	AA977235	Hs.128830 ESTs, Weakly similar to Z192_HUMAN ZINC	1.00	1.00
	442159	AW163390	Hs.278554 heterochromatin-like protein 1	1.92	1.66
25	442179	AA983842	Hs.333555 chromosome 2 open reading frame 2	27.22	50.00
	442328	AI952430	Hs.150614 ESTs, Weakly similar to ALU4_HUMAN ALU S	5.00	3.42
	442432	BE093589	Hs.38178 hypothetical protein FLJ23468	181.59	76.00
	442530	AI580830	Hs.176508 Homo sapiens cDNA FLJ14712 fis, clone NT	10.59	144.00
	442547	AA308997	Hs.217484 ESTs, Weakly similar to ALU1_HUMAN ALU S	109.23	98.00
30	442556	AL137761	Hs.8379 Homo sapiens mRNA; cDNA DKFZp586L2424 (f	1.00	53.00
	442619	AA447492	Hs.20183 ESTs, Weakly similar to AF164793 1 prota	29.02	50.00
	442710	AI015631	Hs.23210 ESTs	1.00	19.00
	442717	R88362	Hs.180591 ESTs, Weakly similar to T23976 hypotheti	1.00	5.00
	442875	BE623003	Hs.23625 Homo sapiens clone TCCCTA00142 mRNA sequ	22.85	50.00
35	442914	AW188551	Hs.99519 hypothetical protein FLJ14007	25.33	82.00
	442932	AA457211	Hs.8858 bromodomain adjacent to zinc finger doma	3.18	4.41
	442942	AW167087	Hs.131562 ESTs	8.45	64.00
	443068	AI188710	ESTs	1.00	27.00
	443204	AW205878	Hs.29643 Homo sapiens cDNA FLJ13103 fis, clone NT	1.00	24.00
40	443211	AI128388	Hs.143655 ESTs	12.42	2.00
	443247	BE614387	Hs.333893 c-Myc target JPO1	128.84	96.00
	443324	R44013	Hs.164225 ESTs	0.02	4.59
	443383	AI792453	Hs.166507 ESTs	1.00	47.00
	443400	R28424	Hs.250648 ESTs	18.52	61.00
45	443426	AF098158	Hs.9329 chromosome 20 open reading frame 1	4.02	1.75
	443572	AA025610	Hs.9605 cleavage and polyadenylation specific fa	2.98	2.57
	443575	AI078022	Hs.269636 ESTs, Weakly similar to ALU1_HUMAN ALU S	1.00	29.00
	443614	AV655386	Hs.7645 fibrinogen, B beta polypeptide	1.00	16.00
	443633	AL031290	Hs.9654 similar to pregnancy-associated plasma p	1.00	39.00
50	443648	AI085377	Hs.143610 ESTs	39.81	70.00
	443715	AI583187	Hs.9700 cyclin E1	48.74	7.00
	443723	AI144442	Hs.157144 syntaxin 6	1.29	1.30
	443802	AW504924	Hs.9805 KIAA1291 protein	1.75	1.61
	443859	NM_013409	Hs.9914 follistatin	1.35	1.13
55	443892	AA401369	Hs.190721 ESTs	1.00	17.00
	443947	W24187	gb:zb47f09.1 Soares fetal lung, NbHL19W	1.33	1.64
	443991	NM_002250	Hs.10082 potassium intermediate/small conductance	5.71	6.87
	444006	BE395085	Hs.10086 type I transmembrane protein Fn14	1.47	1.92
	444009	AI380792	Hs.135104 ESTs	1.00	77.00
60	444017	U04840	Hs.214 neuro-oncological ventral antigen 1	1.00	1.00
	444127	N63620	Hs.13281 ESTs	1.00	29.00
	444129	AW294292	Hs.256212 ESTs	1.00	1.00
	444279	U62432	Hs.89605 cholinergic receptor, nicotinic, alpha p	0.60	7.80
	444371	BE540274	Hs.239 forkhead box M1	2.91	1.14
65	444378	R41339	Hs.12569 ESTs	1.00	1.00
	444381	BE387335	Hs.283713 ESTs, Weakly similar to S64054 hypotheti	469.00	556.00
	444461	R53734	Hs.25978 ESTs, Weakly similar to 2109260A B cell	12.88	105.00
	444471	AB020684	Hs.11217 KIAA0877 protein	24.91	90.00
	444489	AI151010	Hs.157774 ESTs	1.00	111.00
70	444619	BE538082	Hs.8172 ESTs, Moderately similar to A46010 X-in	1.00	70.00
	444665	BE613126	Hs.47783 B aggressive lymphoma gene	30.56	139.00
	444707	AI188613	Hs.41690 desmocollin 3	1.00	1.00
	444735	BE019923	Hs.243122 hypothetical protein FLJ13057 similar to	77.02	90.00
	444781	NM_014400	Hs.11950 GPI-anchored metastasis-associated prote	1.47	1.31
75	444783	AK001468	Hs.62180 anillin (Drosophila Scraps homolog), act	77.55	2.00
	445236	AK001676	Hs.12457 hypothetical protein FLJ10814	1.00	27.00
	445258	AI635931	Hs.147613 ESTs	1.00	73.00
	445413	AA151342	Hs.12677 CGI-147 protein	28.14	50.00
	445417	AK001058	Hs.12680 Homo sapiens cDNA FLJ10196 fis, clone HE	1.81	2.62
80	445443	AV653838	Hs.322971 ESTs	1.00	1.00
	445462	AA378776	Hs.288649 hypothetical protein MGC3077	2.09	1.70
	445517	AF208855	Hs.12830 hypothetical protein	1.87	70.00
	445537	AJ245671	Hs.12844 EGF-like domain, multiple 6	1.71	2.72
	445580	AF167572	Hs.12912 skbl (S. pombe) homolog	1.52	1.34
85	445654	X91247	Hs.13046 thioredoxin reductase 1	1.51	1.52

	445669	AI570830	Hs.174870	ESTs	10.95	11.45
	445818	BE045321	Hs.136017	ESTs	1.00	1.00
	445873	AA250970	Hs.251946	poly(A)-binding protein, cytoplasmic 14	49.42	54.00
	445885	AI734009	Hs.127699	KIAA1603 protein	1.00	132.00
5	445898	AF070623	Hs.13423	Homo sapiens clone 24468 mRNA sequence	1.00	1.00
	445903	AI347487	Hs.132781	class I cytokine receptor	1.00	36.00
	445932	BE046441	Hs.333555	Homo sapiens clone 24859 mRNA sequence	2.41	2.88
	445982	BE410233	Hs.13501	pescadillo (zebrafish) homolog 1, contai	1.60	1.35
	446078	AI339982	Hs.156061	ESTs	1.00	42.00
10	446102	AW168067	Hs.317694	ESTs	1.00	1.00
	446157	BE270828	Hs.131740	Homo sapiens cDNA: FLJ22562 fis, clone H	1.70	1.53
	446269	AW263155	Hs.14559	hypothetical protein FLJ10540	73.01	48.00
	446292	AF081497	Hs.279682	Rh type C glycoprotein	1.55	1.26
	446293	AA420213	Hs.149722	ESTs	1.00	2.00
15	446423	AW139655	Hs.150120	ESTs	1.10	4.19
	446428	AW082270	Hs.12496	ESTs, Weakly similar to ALU4_HUMAN ALU S	0.53	3.26
	446432	AI377320	Hs.150058	ESTs	1.00	5.00
	446528	AU076640	Hs.15243	nucleolar protein 1 (120kD)	1.36	1.31
20	446574	AI310135	Hs.335933	ESTs	3.89	72.00
	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	32.03	20.23
	446636	AC002563	Hs.15767	citron (rho-interacting, serine/threonin	4.19	5.07
	446783	AW138343	Hs.141867	ESTs	2.82	9.47
	446839	BE091926	Hs.16244	mitotic spindle coiled-coil related prot	110.28	28.00
	446849	AU076617	Hs.16251	cleavage and polyadenylation specific fa	3.26	2.94
25	446856	AI814373	Hs.164175	ESTs	6.38	11.30
	446872	X97058	Hs.16362	pyrimidinergic receptor P2Y, G-protein c	1.98	2.03
	446880	AI811807	Hs.108646	Homo sapiens cDNA FLJ14934 fis, clone PL	94.90	113.00
	446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	1.67	3.90
	446989	AK001898	Hs.16740	hypothetical protein FLJ11036	2.82	3.12
30	447022	AW291223	Hs.157573	ESTs	1.00	170.00
	447033	AI357412	Hs.157601	ESTs	7.15	107.00
	447078	AW885727	Hs.9914	ESTs	47.24	24.00
	447081	Y13896	Hs.17287	potassium inwardly-rectifying channel, s	0.12	17.88
	447131	NM_004585	Hs.17466	retinoid acid receptor responder (tazaro	0.97	1.48
35	447149	BE299857	Hs.326	TAR (HIV) RNA-binding protein 2	1.24	1.26
	447153	AA805202	Hs.315562	ESTs	1.00	54.00
	447164	AF026941	Hs.17518	Homo sapiens c1g5 mRNA, partial sequence	1.00	67.00
	447178	AW594641	Hs.192417	ESTs	3.42	50.00
40	447250	AI878909	Hs.17883	protein phosphatase 1G (formerly 2C), ma	1.60	1.52
	447289	AW247017	Hs.36978	melanoma antigen, family A, 3	1.00	1.00
	447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	28.63	1.00
	447343	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m	146.62	51.00
	447350	AI375572	Hs.172634	ESTs	1.00	12.00
	447377	N27687	Hs.334334	transcription factor AP-2 alpha (activat	2.55	63.00
45	447415	AW937335	Hs.28149	ESTs, Weakly similar to KF38_HUMAN KINES	0.91	1.13
	447425	AI963747	Hs.18573	acylphosphatase 1, erythrocyte (common)	1.00	35.00
	447519	U46258	Hs.339665	ESTs	59.89	49.00
	447532	AK000614	Hs.18791	hypothetical protein FLJ20607	1.23	1.63
50	447534	AA401369	Hs.190721	ESTs	1.00	17.00
	447636	Y10043		high-mobility group (nonhistone chromoso	1.41	1.11
	447688	N87079	Hs.19236	Target CAT	1.00	39.00
	447733	AF157482	Hs.19400	MAD2 (mitotic arrest deficient, yeast, h	1.17	1.12
	447769	AW873704	Hs.320831	Homo sapiens cDNA FLJ14597 fis, clone NT	6.47	5.95
55	447802	AW593432	Hs.161455	ESTs	0.73	2.34
	447850	AB018298	Hs.19822	SEC24 (S. cerevisiae) related gene famil	86.45	116.00
	447924	AI817226	Hs.313413	ESTs, Weakly similar to T23110 hypotheti	1.00	1.00
	447973	AB011169	Hs.20141	similar to S. cerevisiae SSIM4	3.50	4.27
	448030	N30714	Hs.325960	membrane-spanning 4-domains, subfamily A	4.13	142.00
60	448105	AI538613	Hs.29824	Transmembrane protease, serine 3	1.15	2.24
	448243	AW369771	Hs.52620	integrin, beta 8	15.84	1.00
	448278	W07369	Hs.11782	ESTs	0.97	1.90
	448290	AK002107	Hs.20843	Homo sapiens cDNA FLJ11245 fis, clone PL	1.00	1.00
	448296	BE622756	Hs.10949	Homo sapiens cDNA FLJ14162 fis, clone NT	2.42	2.17
65	448357	BE274396	Hs.108923	RAB38, member RAS oncogene family	1.44	1.08
	448390	AL035414	Hs.21068	hypothetical protein	1.00	43.00
	448469	AW504732	Hs.21275	hypothetical protein FLJ11011	2.63	2.49
	448569	BE382657	Hs.21486	signal transducer and activator of trans	1.84	2.53
	448663	BE614599	Hs.106823	hypothetical protein MGC14797	3.29	46.00
70	448672	AI955511	Hs.225106	ESTs	1.00	21.00
	448733	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte	1.82	1.08
	448741	BE614564	Hs.19574	hypothetical protein MGC5469	2.48	1.92
	448757	AI365784	Hs.48820	TATA box binding protein (TBP)-associate	23.53	20.00
	448775	AB025237	Hs.388	nudix (nucleoside diphosphate linked moi	2.34	1.97
75	448826	AI580252	Hs.293246	ESTs, Weakly similar to putative p150 [H	74.07	62.67
	448830	AL031658	Hs.22181	hypothetical protein DJ310013.3	1.37	1.31
	448844	AI581519	Hs.177164	ESTs	1.00	31.00
	448988	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	1.84	1.95
	448993	AI471630		KIAA0144 gene product	1.63	1.49
80	449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	1.00	1.00
	449029	N28989	Hs.22891	solute carrier family 7 (cationic amino	1.97	2.26
	449040	AF040704	Hs.149443	putative tumor suppressor	0.97	1.56
	449048	Z45051	Hs.22920	similar to S69401 (cattle) glucose induc	27.13	90.00
	449053	AI625777	Hs.344766	ESTs	8.33	44.00
	449054	AF148848	Hs.22934	myoneurin	73.85	104.00
85	449101	AA205847	Hs.23016	G protein-coupled receptor	2.58	27.00

	449167	T05095	Hs.19597	KIAA1694 protein	1.61	2.35
	449207	AL044222	Hs.23255	nucleoporin 155kD	2.36	1.55
	449228	AJ403107	Hs.148590	protein related with psoriasis	1.15	1.15
5	449230	BE613348	Hs.211579	metastoma cell adhesion molecule	206.65	151.00
	449305	AJ638293		gbttd09b07.x1 NCL CGAP_GC6 Homo sapiens	17.28	45.00
	449318	AW235021	Hs.78531	Homo sapiens, Similar to RIKEN cDNA 5730	26.39	35.00
	449448	D60730	Hs.57471	ESTs	1.00	1.00
	449457	AW205006	Hs.197042	ESTs	1.00	1.00
10	449523	NM_000579	Hs.54443	chemokine (C-C motif) receptor 5	56.80	216.86
	449722	BE280074	Hs.23960	cyclin B1	150.03	1.00
	449976	H06350	Hs.135056	Human DNA sequence from clone RPS-650E9	2.16	2.85
	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitter)	1.17	1.45
	450098	W27249	Hs.8109	hypothetical protein FLJ21080	1.79	2.38
	450101	AV649989	Hs.24385	Human hbc547 mRNA sequence	1.00	69.00
15	450149	AW969781	Hs.132863	Zic family member 2 (odd-paired Drosophila)	1.00	1.00
	450193	AJ916071	Hs.15607	Homo sapiens Fanconi anemia complementat	29.85	34.00
	450221	AA328102	Hs.24641	cytoskeleton associated protein 2	1.00	1.00
	450372	BE218107	Hs.202436	ESTs	1.00	1.00
20	450375	AA009647	Hs.8850	a disintegrin and metalloproteinase doma	51.26	93.00
	450447	AF212223	Hs.25010	hypothetical protein P15-2	123.20	181.00
	450558	AL050078	Hs.25159	Homo sapiens cDNA FLJ10784 fis, clone NT	1.00	19.00
	450589	AJ701505	Hs.202526	ESTs	1.00	23.00
	450684	AA872605	Hs.25333	Interleukin 1 receptor, type II	1.00	100.00
25	450701	H39960	Hs.288467	Homo sapiens cDNA FLJ12280 fis, clone MA	1.89	1.55
	450705	U90304	Hs.25351	iroquois homeobox protein 2A (IRX-2A) (1.00	45.00
	450832	AA401369	Hs.190721	ESTs	25.17	17.00
	450937	R49131	Hs.26267	ATP-dependant interferon response protei	90.92	90.00
	450983	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	3.33	1.70
30	451105	AJ761324		gbwif06b11.x1 NCL CGAP Co16 Homo sapiens	15.02	124.00
	451110	AJ955040	Hs.265398	ESTs, Weakly similar to transformation-r	1.00	143.00
	451253	H48299	Hs.26126	claudin 10	3.02	2.29
	451291	R39288	Hs.6702	ESTs	1.00	1.00
	451320	AW498974		diacylglycerol kinase, zeta (104kD)	2.92	18.00
35	451380	H09280	Hs.13234	ESTs	6.90	6.67
	451386	AB029006	Hs.26334	spastic paraplegia 4 (autosomal dominant	35.75	72.00
	451437	H24143	Hs.31945	hypothetical protein FLJ11071	1.00	69.00
	451462	AK000367	Hs.26434	hypothetical protein FLJ20360	1.83	2.10
	451524	AK001466	Hs.26516	hypothetical protein FLJ10604	1.13	1.07
40	451541	BE279383	Hs.26557	plakophilin 3	1.88	1.33
	451592	AJ805416	Hs.213897	ESTs	1.00	1.00
	451635	AA018899	Hs.127179	cryptic gene	1.52	1.92
	451743	AA401369	Hs.190721	ESTs	4.95	17.00
	451806	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	13.55	31.00
45	451807	W52854		hypothetical protein FLJ23293 similar to	1.55	35.00
	451871	AJ821005	Hs.118599	ESTs	1.81	2.53
	451952	AL120173	Hs.301663	ESTs	1.00	22.00
	452012	AA307703	Hs.279766	kinesin family member 4A	3.43	2.26
	452046	AB018345	Hs.27657	KIAA0802 protein	56.59	19.00
50	452194	AJ694413	Hs.332649	olfactory receptor, family 2, subfamily	1.67	4.09
	452206	AW340281	Hs.33074	Homo sapiens, clone IMAGE:3606519, mRNA,	9.31	53.00
	452240	AA401369	Hs.190721	ESTs	13.42	17.00
	452256	AK000933	Hs.28661	Homo sapiens cDNA FLJ10071 fis, clone HE	39.03	94.00
	452281	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fis, clone PL	153.01	340.00
55	452291	AF015592	Hs.28853	CDC7 (cell division cycle 7, S. cerevisi	1.95	23.00
	452295	BE379336	Hs.28866	programmed cell death 10	42.33	61.00
	452304	AA025386	Hs.61311	ESTs, Weakly similar to S10590 cysteine	1.17	2.14
	452340	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	1.00	13.00
60	452349	AB026944	Hs.29189	ATPase, Class VI, type 11A	1.09	1.42
	452367	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	54.49	53.00
	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro	1.00	32.00
	452410	AL133619		Homo sapiens mRNA; cDNA DKFZp434E2321 (1.26	1.99
	452461	N78223	Hs.108106	transcription factor	24.47	35.00
	452571	W31518	Hs.34665	ESTs	54.61	102.00
65	452613	AA461599	Hs.23459	ESTs	1.39	1.32
	452699	AW295390	Hs.213062	ESTs	1.00	26.00
	452705	H49805	Hs.246005	ESTs	1.00	1.00
	452747	AF160477	Hs.61460	Ig superfamily receptor LNIR	112.87	1.29
	452787	AW294022	Hs.222707	KIAA1718 protein	1.00	1.00
70	452795	AW392555	Hs.18878	hypothetical protein FLJ21620	1.00	1.00
	452823	AB012124	Hs.30696	transcription factor-like 5 (basic helix	7.91	75.00
	452833	BE559691	Hs.30736	KIAA0124 protein	3.16	1.92
	452838	U65011	Hs.30743	preferentially expressed antigen in mela	174.35	1.00
	452862	AA401369	Hs.190721	ESTs	98.26	17.00
75	452865	AW173720	Hs.345805	ESTs, Weakly similar to A47582 B-cell gr	1.55	1.00
	452934	AA581322	Hs.4213	hypothetical protein MGC16207	1.73	1.19
	452946	X95425	Hs.31092	EphA5	1.00	1.00
	452976	R44214	Hs.101189	ESTs	1.58	1.98
	453028	AB005532	Hs.31442	RecQ protein-like 4	1.60	1.60
80	453095	AW295660	Hs.252756	ESTs	0.77	1.50
	453102	NM_007197	Hs.31664	frizzled (Drosophila) homolog 10	1.00	1.00
	453103	AJ301052	Hs.153444	ESTs	1.00	1.00
	453120	AA292891	Hs.31773	pregnancy-induced growth inhibitor	1.23	1.20
	453153	N53893	Hs.24360	ESTs	1.00	83.00
	453160	AJ263307	Hs.239884	H2B histone family, member L	1.00	30.00
85	453197	AW16269	Hs.109057	ESTs, Weakly similar to ALU5_HUMAN ALU S	1.00	134.00

5	453210	AL133161	Hs.32360	hypothetical protein FLJ10857	1.69	1.93		
	453240	A1969564	Hs.166254	hypothetical protein DKFZp566i133	1.00	1.00		
	453317	NM_002277	Hs.41696	keratin, hair, acidic, 1	1.19	1.27		
	453323	AF034102	Hs.32951	solute carrier family 29 (nucleoside tra	4.90	4.11		
	453331	A1240665	Hs.8850	ESTs	199.42	340.00		
	453392	U23752	Hs.32964	SRY (sex determining region Y)-box 11	1.00	16.00		
	453431	AF094754	Hs.32973	glycine receptor, beta	1.00	1.00		
	453439	A1572438	Hs.32976	guanine nucleotide binding protein 4	3.44	5.17		
	453459	BE047032	Hs.257789	ESTs	2.84	5.58		
10	453563	AW608906.comp	Hs.181163	hypothetical protein MGC5629	4.58	90.00		
	453633	AA357001	Hs.34045	hypothetical protein FLJ20764	1.74	1.60		
	453775	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	19.49	1.00		
	453830	AA534296	Hs.20963	ESTs	24.92	25.00		
	453857	AL080235	Hs.35861	DKFZP566E1621 protein	167.59	66.00		
15	453867	A1929383	Hs.33032	hypothetical protein DKFZp434N185	1.00	39.00		
	453883	A1638516	Hs.347524	cofactor required for Sp1 transcriptiona	1.97	1.58		
	453884	AA355925	Hs.36232	KIAA0186 gene product	63.89	20.00		
	453900	AW003582	Hs.226414	ESTs, Weakly similar to ALU8_HUMAN ALU S	20.41	16.00		
20	453922	AF053306	Hs.36708	budding uninhibited by benzimidazoles 1	7.09	22.00		
	453941	U39817	Hs.36820	Bloom syndrome	29.75	19.00		
	453964	A1961486	Hs.12744	ESTs	1.00	1.00		
	453968	AA847843	Hs.62711	Homo sapiens, clone IMAGE:3351295, mRNA	2.06	1.81		
	453976	BE463830	Hs.163714	ESTs	3.02	131.00		
25	454024	AA933527	Hs.293907	hypothetical protein FLJ23403	1.00	131.00		
	454034	NM_000691	Hs.575	aldehyde dehydrogenase 3 family, member	1.23	1.02		
	454042	T19228	Hs.172572	hypothetical protein FLJ20093	30.63	171.00		
	454059	NM_003154	Hs.37048	slatherin	1.00	1.00		
	454066	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid	1.01	1.45		
30	454098	W27953	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	1.26	1.11		
	454241	BE144666	Hs.110826	gb:CM2-HT0176-041099-017-c02 HT0176 Homo	6.33	5.04		
	454417	A1244459	Hs.110826	trinucleotide repeat containing 9	4.30	7.82		
	454439	AW819152	Hs.154320	DKFZP566O1646 protein	1.00	1.00		
	455175	AW993247	Hs.816	gb:RC2-BN0033-180200-014-h09 BN0033 Homo	13.75	103.00		
35	455601	A1368680	Hs.816	SRY (sex determining region Y)-box 2	206.11	1.00		
	456237	AA203682	Hs.816	gb:zx52e07.r1 Soares_fetal_liver_spleen_	1.00	1.00		
	456321	NM_001327	Hs.87225	cancer/testis antigen	1.14	1.10		
	456475	NM_000144	Hs.95998	Friedreich ataxia	1.00	48.00		
	456508	AA502764	Hs.123469	ESTs, Weakly similar to AF208855 1 BM-01	162.25	189.00		
40	456534	X91195	Hs.100623	phospholipase C, beta 3, neighbor pseudo	2.12	1.80		
	456736	AW248217	Hs.1619	achaete-scute complex (Drosophila) homol	1.15	1.94		
	456759	BE259150	Hs.127792	delta (Drosophila)-like 3	1.00	1.00		
	456990	NM_004504	Hs.171545	HIV-1 Rev binding protein	16.42	84.00		
	457200	U33749	Hs.197764	thyroid transcription factor 1	0.57	1.76		
45	457234	AW968360	Hs.14355	Homo sapiens cDNA FLJ13207 fis, clone NT	2.71	4.15		
	457465	AW301344	Hs.122908	DNA replication factor	46.37	47.00		
	457489	A1693815	Hs.127179	cryptic gene	1.12	1.35		
	457646	AA725650	Hs.112948	ESTs	1.55	2.51		
	457733	AW974812	Hs.291971	ESTs	1.00	55.00		
	457819	AA057484	Hs.35406	ESTs, Highly similar to unnamed protein	4.36	3.18		
50	458092	BE545684	Hs.343566	KIAA0251 protein	1.00	1.32		
	458098	BE550224	Hs.7655	metallothionein 1E (functional)	1.00	22.00		
	458207	T28472	Hs.28465	U2 small nuclear ribonucleoprotein auxil	2.06	1.88		
	458242	BE299588	Hs.28465	Homo sapiens cDNA: FLJ21869 fis, clone H	1.00	1.00		
55	458247	R14439	Hs.209194	ESTs	7.00	9.85		
	458679	AW975460	Hs.142913	ESTs	1.00	3.00		
	458778	AW451034	Hs.326525	arylsulfatase D	1.31	2.01		
	458933	A1638429	Hs.24783	RAN binding protein 1	1.98	1.71		
	459352	AW810383	Hs.206828	ESTs	12.60	63.00		
60	459670	F01020	Hs.172004	titin	1.00	1.00		
	459702	A1204995	Hs.172004	gb:an03c03.x1 Stratagene schizo brain S1	1.00	237.00		

TABLE 9B

65	Pkey:	Unique Eos probeset identifier number
	CAT number:	Gene cluster number
	Accession:	Genbank accession numbers
70	Pkey	CAT Number
	407746	10125_1
		AK001962 R69415 BE464605 AA418699 AA053293 AA149075 AA058396 AW338226 AW272659 AA454607 A1139535 AW469852 A1275461
		AW271982 AA730033 AA576507 AA991217 AA782067 A1985851 AA805864 AA505598 AW469857 R69545 AA988279 AW001647 N63320
		D82661 T27343 AA306950 AA360989 R58778
	408070	1036688_1
		AW148852 BE350895
75	408650	107294_1
		AA525775 AA056342 A1538978 AW975281 AA664986
	409522	113735_1
		AA075382 AA075431
	409856	1156522_1
		AW502152 H41202 H29772
	410032	1170435_1
		BE065985 BE065944 BE066008 BE066083 BE066093
	411089	123172_1
		AA456454 AA713730 AA091294 AA584921 N86077 AW836781 AA601031 AA579876 AA551106 AA633188 AW905577 A1955808 A1679386
		A1679895 AA514764 AA454562 A1082382 AA595822 AA551351 AA586369 AA666384 AA188934 AA666398 AA551297 AA565183
80	411152	1234028_1
		BE069199 AW936012 AW877466 AW819782 AW935798 AW835546 AW936042 BE069121 AW835625 AW877536 AW935885 BE069202
		AW820019 AW935937 BE160180 AW935946 BE069101 BE069125 AW877527 BE160316 BE160399 AW935794 AW835701 AW935784
	412537	1304_1
		AL031778 X59711 NM_002505 M59079 A1870439 A194259 AW664010 AA405063 AA436132 BE174516 AA412691 A1400314 AA436024
		T29403 BE079412 BE079428 N90322 A1631202 A1141758 A1016793 A167566 A1862075 A1375230 A1209445 AW235763 A1044113 AA382556
		AW953918 AA927051 AA889823 BE003094 AW390155 AW360805 AW360823 AW360810 AA425472 A1694282 A1044114 A1684577 A1609865

WO 02/086443

PCT/US02/12476

			A478773 A160445 A1674630 N69088 AW665529 N49278 A1129239 A457890 A1621264 AW297152 A1268215 AA907787 A1286170 A1017982 A1863541 A1469807 A1969353 BE552356 N66509 A4736741 A4382555 AW075811 AW292026 H06382 AW957730 AA352014 R13591 A1121201 D60420 BE263253 BE047862 Z41952 A424991 A1693507 A1863108 AA599060 A1091148 AA598689 R39887 AA813482 AW016452 H06383 R41807 A1364268 AA620528 A1241940 AW089149 AW090733 AW088575 Z38240 AA121202 R17734 BE157489 BE157560 AA926960 AA926959 W76521 W24270 W21526 AA037172 BE267636 H83186 AA469909 N86396 AA001348 BE535736 AA081745 BE566245 AA082436 H72525 H77575 N49786 W80565 H78746 BE569085 W04339 R98127 T55938 BE279271 AW96304 T29812 AA476873 BE297387 AA292753 AA177048 NM_001826 X54941 BE314366 AA908783 A1719075 BE270172 BE269819 AA889955 A1204530 W25243 A1935150 AA872039 W72295 T99630 A142691 H98460 N31428 BE255916 H03265 A1857576 AA776920 AA910644 AA459522 AA293140 AW514667 R75953 AW662396 AA662522 A1865147 A1423153 AW262230 AA584410 AA583187 AW024595 AW069734 A1826996 AA282997 AA876046 AW613002 AA527373 AW972459 A1831360 AA621337 AA100926 AA772418 AA594628 A1033892 W95096 A1034317 AA398727 A1085031 N95210 A1459432 A1041437 AA932124 AA627684 AA935829 A1004827 A1423513 A1094597 H42079 R54703 A1530359 AA617681 AA978045 AA643280 W44561 A191988 A1537692 A1090262 AA740817 A1312104 A1911822 AA416871 A1185409 AA129784 AA701623 A1075239 A1139549 AA633648 A1339996 A1336880 AA399239 A1078708 A1085351 A1362835 A1346618 A1146955 A1989380 A1348243 N92892 AA765850 A1494230 A1278887 AA962596 A1492600 W80435 AA001979 R97424 A125015 N24127 AA157451 AA235549 AA459292 AA037114 AA129785 A1494211 AW059601 AW886710 R92790 N59755 A1361128 AW589407 H47725 H97534 H48076 H48450 T99631 AW300758 H03431 R76789 AA954344 H77576 R96823 A1457100 N92845 N49682 H42038 BE220698 BE220715 H99552 AA701624 N74173 R54704 H79520 H72923 H03266 BE261919 AA769633 AA480310 AA507454 AA910586 A1203723 AW104725 W25611 W25071 T88980 H03513 T77589 R99156 W95095 R97470 AA702275 T77551 AA911952 H82956 N83673 AA283672 A1267700 A1720344 AA191424 A1023543 A1469633 AA172056 AW958465 AA172236 AW953397 AA355086 AW265494 AA455904 AA195677 AW265432 AW991605 AA456370 N28754 N28747 A1568146 A1979339 AA322671 AA322672 AW955043 A1990326 AA776406 A1016250 AA843578 AW451882 N23137 N23129 W70051 A1038748 AA831327 A1925845 AW945895 Z42183 T31621 T97478 D62703 AA242966 D79798 A1076704 T74854 T74860 T72098 T73265 T73873 T69180 T74658 T58786 T60385 T73410 T68781 T67845 T67593 T73952 T67864 T60630 T68367 T68401 T53959 T72360 T72099 T60377 T56961 T71712 T72821 T64738 T74645 T72037 T68688 T72063 T73258 T72326 T64242 T68220 T74673 T71800 T68355 T61227 T62738 T69317 T53850 T64692 T73768 T73962 T73382 T69914 T70975 T73400 T60631 T73277 T73203 T70498 T61409 T58925 NM_000508 M64982 T68301 T73729 T69445 T60424 T67922 T67736 T68716 T67755 T74765 T73819 T58719 T74756 T60477 T74863 T61109 T68329 T58850 T71857 T73425 T53736 T66607 T58898 T64309 T72031 T72079 T64305 T71908 T68107 T71916 T73787 T56035 T64425 T71870 T60476 T61376 T67820 T71895 T41006 T69441 T68170 T74617 T71958 T69440 T61875 R06796 H48353 T71914 T53939 T64121 AA693996 T72525 T67779 T68078 AA011465 AA345378 AV654847 AV654272 AV656001 A1064740 T82897 N33594 AA344542 AW805054 A1207457 T61743 AA026737 H94389 AA382695 AA918409 T68044 S82092 T39959 A1017721 AA312395 AA312919 T40156 H66239 AV652989 H38728 R98521 AV655200 R95790 W03250 W00913 AA344136 AV660126 R97923 AA343596 AW470774 AV651256 N54417 AA812862 AW182929 A111192 H61463 H72060 AA344503 H38639 A1277511 AV651108 A1027625 T47810 AA235252 T27853 T47778 R95746 H70620 AA701453 AW827166 R98475 C20925 AV657287 T71959 T71313 T73320 T73333 T61618 T69293 T69283 T73931 T72178 T72456 AV645639 AV653476 T72957 T72300 T58906 T71457 T70494 T72956 T70495 T68267 T74407 T85778 AA344726 T27854 T74485 T74101 T73868 T71518 T72304 AA343853 T73909 T68070 T72065 H72149 T73493 T73495 AV645993 R02293 T70475 T64751 AA344441 AA343557 AA345732 AA344328 A1110639 AA344603 AF063513 T64696 T68516 T72223 T60507 T67633 R29500 T72517 R02292 T60599 T69206 T70452 T74677 R29366 T61277 T74914 T60352 R29675 T74843 AV645792 AA344408 T69197 T72057 T69368 T69358 T68258 AV650429 T73341 T61702 T74598 T40095 K02272 T40106 AA343045 AA341908 AA342807 AA341964 T53747 T72042 T62764 A1064899 AA343060 T67832 T72440 T71770 T68091 T69108 T72449 T69167 T71289 T68251 AV654844 T64375 AA345234 T67598 AA011414 T68036 H48262 A1207557 T68219 W86031 T69081 T64232 R93196 T62136 AV650539 H67459 T72978 AA344583 T60362 H58121 T95711 T72803 T68055 T71715 R29036 T72793 T69122 T64595 T62888 T69139 T68291 T64652 T67971 T46862 AA693592 A1248502 R29454 T64764 T57001 T73052 T71429 T51176 T58866 AV655414 H90426 AA342489 T73666 T67848 T72512 T53835 T67837 T73317 T74273 T69420 T68245 T74380 T67862 T74474 T56068 A1972768 BE142230 AA252019 A1910275 X00474 X52003 X05030 NM_003225 AA314326 AA308400 AA506787 AA314825 A1571948 AA507595 AA614579 AA587613 R83818 AA568312 AA614409 AA307578 A1925552 AW950155 A1910883 M12075 BE074052 AW004668 AA578674 AA582094 BE074053 BE074126 BE074140 AA514776 AA588034 BE074051 BE074068 AW009769 AW050690 AA582876 R55389 A1001051 AW050700 AW750216 AA614539 BE074045 A1307407 AW602303 BE073575 A1202532 AA524242 A1970839 A1909751 BE076078 A1909749 R55292 AW881145 AA490718 M85637 AA304575 T06067 AA331991 AL119930 AA320696 AW752565 AL031985 AL137241 A1792386 A1733654 A1857654 A1049911 AA337221 AA336756 AW966196 AW953120 R56325 AA349562 A1493134 A1498691 AW771508 A1498457 A1768408 A1783624 A1383985 A1580267 D79813 AA393768 AK001536 AA191092 AW510354 A1554256 AL353968 AA134266 AA663848 AA400100 AA401424 AL038943 AA161338 BE268213 AA425597 N87306 AA092959 BE566038 AA247451 N47392 A1928802 AW182594 AW027872 A1819831 A1936994 W56258 A1653448 A1278611 A1283557 A1824306 AW338658 AW150899 AA687514 N47393 N29885 AA973469 A1038904 A120264 A1034339 AW674583 N72156 A1079733 A1038683 A1291616 AA491599 AA93675 A1837380 BE006554 BE006473 A1087090 T33044 AA652043 A1203503 AA583959 W35283 A1129926 Z41844 AW020925 AW575848 A1684603 AA493297 A1140689 A1277175 AA425444 A1932767 W02632 BE396786 R37261 AW207206 AW341473 AA448195 A1951341 AA249027 AL038984 AK001993 AL080066 AV652725 BE566226 AA345557 AA315222 AA090585 AA375688 AA301092 AA298454 W05762 AW607939 H51658 D83880 N84323 BE295821 AW947007 D61461 AW079261 AA329482 AW901780 A1354442 AA772275 R31663 A1354441 A1767525 H92431 A1916735 H93575 A1394255 AW014741 A1573090 C06195 AW612857 AW265195 A1339558 A1377532 A1308821 A1919424 A1589705 AW055215 A1336532 A1338051 AA806547 C75509 C00618 AW071172 AW769904 AA630381 A1878018 A1863985 D79662 BE221049 AW265018 A1589700 AW196655 N76573 A1370908 BE042393 N75017 A1698870 AW960115 AL133561 AL041090 AL117481 AL122069 AW439292 A1968826 AW072916 A1184913 AA489195 AW466994 AW469044 N59350 A1819642 A1280239 A1220572 AA789302 A1473611 AW841126 D60937 BE041395 AA491826 AA621946 AA715980 AA665102 AW970622 AA503009 AA502988 AA502805 T92188 AA221036 R87170 BE537068 BE544757 C18935 AW812058 T92565 AA227415 AA233942 AA223237 AA658403 AA601627 AW866939 BE061833 BE000620 AW961170 AW847519 AA308542 AW821833 AW945698 C04699 AA205504 AA377241 AW821667 AA055720 AW817981 AW856468 AA155719 AA179928 T03007 AW754298 AA227407 AA113928 AA307904 C16859 A1798376 S46400 AW811617 AW811616 W00557 BE142245 AW858232 AW861851 AW858362 AA232351 AA218557 AA055556 AW858231 AW857541 AW814172 H66214 AW814398 AF134164 AA243093 AA173345 AA199942 AA223384 AA227092 AA227080 T12379 AA092174 T61138 AA149776 AA699829 AW879188 AW813567 AW813538 A1267168 AA157718 AA157719 AA100472 AA100774 AA130756 AA157705 AA157730 AA157715 AA053524 AW849581 AW854566 C05254 AW882836 T92637 AW812621 AA206583 AA209204 BE156909 AA226824 A1829309 AW991957 N66951 AA527374 H66215 AA045564 A1694265 H60808 AA149726 AW195620 BE081333 BE073424 AW817662 AW817705 AW817703 AW817659 BE081531 H59570 AA626980 A1126603 BE504035
5	412811	132943_1	
	413690	1383256_1	
	414883	15024_1	
10			
15			
20	415989	156454_1	
	417324	166714_1	
	418574	17690_1	
25	418712	1784125_1	
	419443	184788_1	
	419502	18535_1	
30			
35			
40			
45			
50	419936	189181_1	
	421582	2041_1	
55	422128	211994_1	
	423034	224122_1	
	423816	23234_1	
	424200	236595_1	
	424999	245835_1	
	426956	273896_1	
	426991	27415_1	
60	427260	276598_1	
	428023	28589_2	
65			
	429220	301384_1	
	429978	31150_1	
70			
	430439	31808_1	
	430935	325772_1	
	431089	327825_1	
	431322	331543_1	
	432407	34624_1	
75			
80	434414	38585_1	
85	436608	42361_3	

WO 02/086443

PCT/US02/12476

438091	44964_1	AW373062 T56662 A1299190 BE174210 AW579001 H01811 W40186 R67100 A1923886 AW952164 AA628440 AW898607 AW898616 AA709126 AW898628 AW898644 AA947932 AW898625 AW898622 A1276125 A185720 AW510698 AA987230 T52522 BE467708 AW243400 AW043542 A1288245 A1186932 D52654 D55017 D52715 D52477 D53333 D54679 A1298739 A1146984 A1922204 N83843 BE174213 AA845571 A1813854 A1214518 A1635262 A1139455 A1707807 A1698085 AW884528 A1024768 A1004723 AW087420 A1555133 N94964 A1263939 AW513280 A1061126 A1435818 A1859106 A1360506 A1024767 AA513019 AA757598 X56196 AA902959 A1334784 A1860794 AA010207 AW890091 AW513771 A1951391 A1337671 T52499 AA890205 A1640908 H75966 AA463487 AA358688 A1961767 A1865295 AA780994 A1985913 BE174196 AA029094 AW592159 T55581 N79072 A1611201 AA910812 A1220713 AW149306 A1758412 AA045713 R79750 N76096 AW579121 AA847986 AA829098 AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 A1346341 A1867454 N54784 A1655270 A1421279 AW014882 AA775552 N62351 N59253 AA626243 A1341407 BE175639 AA456968 A1358918 AA457077 AL109688 R23665 R26578 AA570256 AW014761 AA573721 A1473237 A1022165 AA554071 AA127551 N90525 AW973623 AA447991 AA243852 BE328850 A148171 A1359627 A1005068 A1356567 AA232991 AW016855 AA906902 AA233101 AA127550 BE512923 A188710 A1032142 AW078833 N30308 AW675532 A1219028 A1341201 N22181 H95390 W24187 W24194 R17789 Y10043 NM 005342 L05085 ALO34450 BE614226 AW749053 AA379173 AA248230 BE514634 AA334622 R70656 AA367593 AA214649 AA369318 AW957081 R05760 AA039903 A1886597 AW630122 AA906264 AA041527 R01145 A1088688 BE463637 AA398795 A1354883 A1768938 A1669995 A1452952 A1168582 A1189869 A1086670 AW262560 AW613954 AA862839 AA435840 AA670197 A1024032 A1990659 A1990089 N81095 AA847919 AW960150 AA211075 AA044704 AA367594 AW582587 AW858854 AW818630 AW818628 AW818433 AW582595 AA096002 N83992 A1471630 BE540637 BE265481 AW407710 BE513882 BE546739 AA053597 BE140503 BE218514 AW956702 A1656234 A1636283 A1567265 AW340858 BE207794 AA053085 R69173 AA292343 AA454908 AA293504 A1659741 A1927478 AA399460 A1760441 AA345416 BE047245 AA730380 AA394063 AA454833 A1982791 A1567270 A1813332 A1767858 AA427705 D20284 A1221458 BE048537 A1263048 AA345417 AA911497 BE537702 A1638293 AW813561 A1761324 AW880941 AW880937 AW118072 A1631982 T15734 AA224195 A1701458 W20198 F26326 AA890570 N90552 AW071907 A1671352 A1375892 T03517 R88265 A124088 AA224388 A1084316 A1354686 T33652 A1140719 A1720211 T03490 A1372637 T15415 AW205836 AA630394 T03515 T33230 AA017131 AA443303 T33623 A1222556 T33511 T33785 A1419606 D55812 W52854 AL117600 BE208116 BE208432 BE206239 BE082291 AW953423 AA351619 BE180648 BE140560 W60080 AA865478 N90291 AW450652 AW449519 AA993634 A1806539 AA351618 AW449522 A1827626 AA904788 AA380381 AA885045 AA774409 BE003229 Z41756 A1133619 AA468118 AA383064 A1476447 T09430 A1673758 AA524895 A1581345 A1300820 AW498812 AA256162 A1559724 A1685732 AA602400 AA905453 A1204595 AW166541 AA157455 AA156269 AA383652 AA431072 AW592707 A1435410 AW27464 A1215594 AA622747 R74039 N35031 A1804128 AW513621 AA868351 A1026826 A1493388 AA614641 W81604 A1567080 A1214351 AA730140 A1125754 A1200813 A1269603 A155082 A1807095 A1476629 AA505909 A1368449 A1686077 A1582930 AW085038 AA757863 AA730154 A1767072 AA468316 A1734130 A1734138 AA426284 AA433997 A1741241 AW043563 A1732741 A1732734 AA437369 AA425820 AA664048 R74130 BE144666 BE184942 AW238414 BE184946 AW993247 AW861464 AA203682 R11958 BE550224 AA832519 N45402 AW885857 N29245 BE465409 W07677 AW970089 A1299731 AA482971 BE503548 H18151 W79223 AF086393 AA461301 W74510 R34182 A1090689 N46003 BE071550 R28075 AW134982 A1240204 A138906 AW026179 A1572316 BE466182 A1206395 A1276154 A1273269 A1422817 A1371014 A1421274 A1188525 AA939164 BE549810 AW137865 A1694996 BE503841 AA459718 BE327407 BE467534 BE218421 BE467767 AA898054 BE467063 A1797130 BE327781
--------	---------	--

TABLE 9C

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
400512	9796593	Minus	1439-1615
400517	9796688	Minus	49998-50346
400560	9843598	Plus	94182-94323, 97058-97243, 101095-101236, 102824-103005
400664	8118496	Plus	13558-13721, 13942-14090, 14554-14679
400665	8118498	Plus	16879-17023
400666	8118496	Plus	17982-18115, 20297-20456
400749	7331445	Minus	9162-9293
400763	8131616	Minus	35537-35784
401027	7230983	Minus	70407-70554, 71060-71160
401093	8516137	Minus	22335-23166
401203	9743387	Minus	172961-173056, 173868-173928
401212	9858408	Plus	87839-88028
401411	7799787	Minus	144144-144329
401435	8217934	Minus	54508-55233
401464	6682291	Minus	170688-170834
401714	6715702	Plus	96484-96681
401747	9789672	Minus	118596-118816, 119119-119244, 119609-119761, 120422-120990, 130161-130381, 130458-130593, 131097-131258, 131866-131932, 132451-132575, 133580-134011
401760	9929699	Plus	83126-83250, 85320-85540, 94719-95287
401780	7249190	Minus	28397-28617, 28920-29045, 29135-29296, 29411-29557, 29705-29787, 30224-30573
401781	7249190	Minus	83215-83435, 83531-83656, 83740-83901, 84237-84393, 84955-85037, 86290-86814
401785	7249190	Minus	165776-165996, 166189-166314, 166408-166569, 167112-167268, 167387-167469, 168634-168942
401797	6730720	Plus	6973-7118
401961	4581193	Minus	124054-124209
401985	2580474	Plus	61542-61750
401994	4153858	Minus	42904-43124, 43211-43336, 44607-44763, 45199-45281, 46337-46732
402075	8117407	Plus	121907-122035, 122804-122921, 124019-124161, 124455-124610, 125672-126076
402260	3399665	Minus	113765-113910, 115653-115765, 116808-116940
402265	3287673	Plus	21059-21168
402297	6598824	Plus	35279-35405, 35573-35659
402408	9796239	Minus	110326-110491

5	402420	5796339	Plus	129750-129319
	402674	8077108	Minus	39290-39502
	402802	3287156	Minus	53242-53432
	402994	2996643	Minus	4727-4969
	403137	9211494	Minus	92349-92572,92958-93084,93579-93712,93949-94072,94591-94748,95214-95337
	403306	8099945	Plus	127100-127251
	403329	8516120	Plus	95450-95598
	403381	9438267	Minus	25009-26178
10	403478	9958258	Plus	116458-116564
	403485	9966528	Plus	2888-3001,3198-3532,3655-4117
	403627	8569879	Minus	23868-24342
	403715	7239669	Plus	85128-85292
	404044	9558573	Minus	225757-225939
15	404076	9931752	Minus	3848-3967
	404101	8076925	Minus	125742-125997
	404140	9843520	Plus	37761-38147
	404165	9925489	Minus	69025-69128
	404185	4572584	Minus	129171-129327
20	404210	5006246	Plus	169926-170121
	404253	9367202	Minus	55675-56055
	404287	2326514	Plus	53134-53281
	404298	9944263	Minus	73591-73723
	404347	9838195	Plus	74493-74829
25	404440	7528051	Plus	80430-81581
	404721	9856648	Minus	173763-174294
	404794	4826439	Plus	101619-101898
	404854	7143420	Plus	14260-14537
	404877	1519284	Plus	1095-2107
30	404927	7342002	Plus	69690-69663
	404996	6007890	Plus	37999-38145,38652-38998,39727-39872,40557-40674,42351-42450
	405449	7622497	Plus	42236-42570
	405568	6006906	Plus	35912-36065
	405572	3800891	Plus	85230-85938
35	405846	4914350	Plus	741-969
	405876	4557087	Plus	73195-73917
	405770	2735037	Plus	61057-62075
	405932	7767812	Minus	123525-123713
	406137	9166422	Minus	30487-31058
40	406360	9256107	Minus	7513-7673
	406399	9256288	Minus	63448-63554
	406467	9795551	Plus	182212-182958

45 TABLE 10A: Potential Therapeutic, Diagnostic and Prognostic targets for Therapy of Lung Cancer and Non-malignant Lung Disease
Table 2A shows about 307 genes up-regulated in non-malignant lung disease relative to lung tumors and normal body tissues and/or down-regulated in lung tumors relative to normal lung and non-malignant lung disease. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

50 Table 10B show the accession numbers for those Pkey's lacking UnigeneID's for table 10A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

55 Table 10C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 10A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

	Pkey:	Unique Eos probeset identifier number				
	ExAccn:	Exemplar Accession number, Genbank accession number				
	UnigeneID:	Unigene number				
60	Unigene Title:	Unigene gene title				
	R1:	Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples				
	R2:	Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples				
65	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
	404394			ENSP00000241075:TRRAP PROTEIN.	0.79	3.10
	404916			Target Exon	1.00	159.00
	405257			Target Exon	1.00	422.00
70	407228	M25079	Hs.155376	hemoglobin, beta	0.47	2.33
	407568	AA740964	Hs.62699	ESTs	1.00	123.00
	408562	AI436323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,	1.00	230.00
	409031	AA376836	Hs.76728	ESTs	1.00	128.00
	410434	AF051152	Hs.63668	tol-like receptor 2	39.65	149.00
	410467	AF102546	Hs.63931	dachshund (Drosophila) homolog	1.00	109.00
75	410808	T40326	Hs.167793	ESTs	1.14	13.14
	412351	AL135960	Hs.73828	T-cell acute lymphocytic leukemia 1	0.37	2.27
	412372	R65998	Hs.285243	hypothetical protein FLJ22029	1.00	173.00
	413795	AL040178	Hs.142003	ESTs	0.10	11.90
80	414154	AW205314	Hs.323060	ESTs	0.62	2.09
	414214	D49958	Hs.75819	glycoprotein M6A	0.03	4.55
	414998	NM_002543	Hs.77729	oxidised low density lipoprotein (lectin	0.64	2.97
	415122	D60708	Hs.22245	ESTs	0.07	8.97
	415765	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and	0.67	1.65
	415775	H00747	Hs.29792	ESTs, Weakly similar to I38022 hypothei	0.29	2.64
85	415910	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	1.00	145.00

	416319	AD815601	Hs.79197	CD83 antigen (activated B lymphocytes, i	15.32	237.00
	416402	NM_000715	Hs.1012	complement component 4-binding protein,	0.64	4.00
	417355	D13168	Hs.82002	endothelin receptor type B	0.01	3.90
5	417421	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m	36.30	357.00
	417511	AL049176	Hs.82223	chordin-like	1.00	179.00
	418489	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h	0.02	6.00
	418726	BE241812	Hs.87860	protein tyrosine phosphatase, non-recept	1.00	113.00
	418741	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom	0.44	1.90
	418883	BE387036	Hs.1211	acid phosphatase 5, tartrate resistant	0.96	2.04
10	419086	NM_000216	Hs.89591	Kallmann syndrome 1 sequence	0.62	2.74
	419150	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous	0.03	6.90
	419235	AW470411	Hs.288433	neurotrophin	1.48	5.13
	419407	AW410377	Hs.41502	hypothetical protein FLJ21276	37.55	336.00
	420556	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	0.80	3.65
15	420656	AA279098	Hs.187636	ESTs	1.65	8.07
	420729	AW964897	Hs.290825	ESTs	2.99	25.82
	421177	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f	0.46	1.95
	422060	R20893	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	1.00	156.00
	422426	W79117	Hs.58559	ESTs	0.03	7.44
20	422652	AW957969	Hs.118958	synixin 11	0.14	3.62
	423099	NM_002837	Hs.123641	protein tyrosine phosphatase, receptor t	0.01	3.16
	424433	H04607	Hs.9218	ESTs	0.75	141.75
	424585	AA464840	Hs.131987	ESTs	1.00	167.00
25	424711	NM_005795	Hs.152175	calcitonin receptor-like	0.43	3.01
	424973	X92521	Hs.154057	matrix metalloproteinase 19	0.37	19.45
	425023	AW956889	Hs.154210	endothelial differentiation, sphingolipi	0.14	3.35
	425664	AJ006276	Hs.159003	transient receptor potential channel 6	1.00	94.00
	425998	AJ0076629	Hs.165950	fibroblast growth factor receptor 4	0.68	1.42
	426657	NM_015865	Hs.171731	solute carrier family 14 (urea transport	0.03	3.74
30	426753	T89832	Hs.170278	ESTs	1.00	141.00
	427558	D49493	Hs.2171	growth differentiation factor 10	1.00	117.00
	427983	M17706	Hs.2233	colony stimulating factor 3 (granulocyte	0.75	2.20
	428467	AK002121	Hs.184465	hypothetical protein FLJ11259	0.76	2.25
	428927	AA441837	Hs.90250	ESTs	0.01	3.62
35	429496	AA453800	Hs.192793	ESTs	1.00	138.00
	430458	NM_004673	Hs.241519	angiotensin-like 1	1.00	132.00
	431385	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A	1.00	157.00
	431728	NM_007351	Hs.268107	multimerin	1.00	157.00
	431848	AJ378857	Hs.126758	ESTs, Highly similar to AF175283 1 zinc	0.34	2.24
40	432128	AA127221	Hs.117037	ESTs	0.00	1.15
	432519	AJ221311	Hs.130704	ESTs, Weakly similar to BCHUIA S-100 pro	0.01	2.06
	433043	W57554	Hs.125019	lymphoid nuclear protein (LAF-4) mRNA	1.00	267.00
	433803	AJ823593	Hs.27688	ESTs	1.00	105.00
	434730	AA644669	Hs.193042	ESTs	1.05	3.15
45	435472	AW972330	Hs.283022	triggering receptor expressed on myeloid	0.83	1.94
	436532	AA721522		gb:nv54h12.r1 NCI_CGAP_Ew1 Homo sapiens	1.00	218.00
	437119	AJ379921	Hs.177043	ESTs	1.00	133.00
	437140	AA312799	Hs.283689	activator of CREM in testis	0.67	122.67
50	437211	AA382207	Hs.5509	ecotropic viral integration site 2B	1.00	142.00
	437960	AJ669586	Hs.222194	ESTs	1.00	147.00
	438202	AW169287	Hs.22588	ESTs	1.00	141.00
	438873	AJ302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	0.71	3.66
	438875	AA827640	Hs.189059	ESTs	23.32	370.00
55	441048	AA913488	Hs.192102	ESTs	0.77	8.50
	441188	AW292830	Hs.255609	ESTs	3.43	16.36
	441499	AW296235	Hs.101689	ESTs	1.00	167.00
	444513	AL120214	Hs.7117	glutamate receptor, ionotropic, AMPA 1	1.00	151.00
	444527	NM_005408	Hs.11383	small inducible cytokine subfamily A (Cy	46.47	153.00
	444561	NM_004469	Hs.11392	c-fos induced growth factor (vascular en	0.01	3.08
60	445279	R41900	Hs.22245	ESTs	0.60	141.00
	446017	N98238	Hs.55185	ESTs	0.18	2.39
	446984	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15	0.10	2.16
	446998	N99013	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	0.01	2.53
	447357	AJ375922	Hs.159367	ESTs	0.46	2.64
65	448108	AJ800470	Hs.171941	ESTs	18.05	296.00
	448253	H25899	Hs.201591	ESTs	1.00	141.00
	449275	AW450848	Hs.205457	perlecan	0.56	1.38
	450400	AJ694722	Hs.279744	ESTs	0.88	4.33
70	450696	AJ654223	Hs.16026	hypothetical protein FLJ23191	0.52	2.08
	450726	AW204600	Hs.250505	retinoic acid receptor, alpha	0.79	2.01
	451497	H83294	Hs.284122	Wnt inhibitory factor-1	0.35	2.03
	451533	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	0.13	2.25
	453636	R67837	Hs.168872	ESTs	1.00	116.00
	458332	AJ000341	Hs.220491	ESTs	1.00	192.00
75	459580	AA022888	Hs.176065	ESTs	0.20	2.98
	400269			Eos Control	0.40	2.40
	403421			NM_016369: Homo sapiens claudin 18 (CLDN	0.53	1.77
	407570	Z19002	Hs.37096	zinc finger protein 145 (Kruppel-Box, e	0.01	3.18
80	412295	AW088826	Hs.117176	poly(A)-binding protein, nuclear 1	0.56	1.74
	414517	M24461	Hs.76305	surfactant, pulmonary-associated protein	0.64	1.50
	417204	N81037	Hs.1074	surfactant, pulmonary-associated protein	0.33	1.16
	418307	U70857	Hs.83974	solute carrier family 21 (prostaglandin	0.53	1.55
	418935	T28499	Hs.89485	carbonic anhydrase IV	0.20	1.28
	421502	AF111856	Hs.105039	solute carrier family 34 (sodium phospho	0.78	1.90
85	421798	N74880	Hs.29977	N-acylsphingosine amidohydrolase (acid c	0.59	1.54

	423354	AB011130	Hs.127436	calcium channel, voltage-dependent, alph	0.59	1.55
	423738	AB002134	Hs.132195	airway trypsin-like protease	10.14	51.00
	425211	M18667	Hs.1867	progastricsin (pepsinogen C)	0.35	1.62
5	425438	T62216	Hs.270840	ESTs	0.23	9.45
	426828	NM_000020	Hs.172670	activin A receptor type II-like 1	0.03	1.71
	427019	AA001732	Hs.173233	hypothetical protein FLJ10970	0.01	1.49
	428043	T92248	Hs.2240	uteroglobin	0.42	1.26
	430280	AA361258	Hs.237868	Interleukin 7 receptor	0.46	2.43
10	431433	X65018	Hs.253495	surfactant, pulmonary-associated protein	0.57	1.59
	431723	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	0.29	1.80
	432985	T92363	Hs.178703	ESTs	0.32	2.27
	441835	AB036432	Hs.184	advanced glycosylation end product-speci	0.31	1.51
	442275	AW449467	Hs.54795	ESTs	0.55	1.78
	443709	AI082692	Hs.134662	ESTs	0.00	3.02
15	444325	AW152618	Hs.16757	ESTs	0.32	2.49
	450954	AI904740	Hs.25691	receptor (calcitonin) activity modifying	0.46	1.74
	451558	NM_001089	Hs.26630	ATP-binding cassette, sub-family A (ABC1	0.52	1.87
	453310	X70697	Hs.553	solute carrier family 6 (neurotransmitte	0.00	3.30
	456855	AF035528	Hs.153863	MAD (mothers against decapentaplegic, Dr	0.01	2.31
20	444342	NM_014398	Hs.10887	similar to lysosome-associated membrane	0.66	2.20
	400754			Target Exon	1.00	297.00
	401045			C11001883*gi 5753278 ref NP_033938.1 c	1.00	109.00
	401083			NM_016582*Homo sapiens peptide transp	0.89	1.39
	402474			NM_004079:Homo sapiens cathepsin S (CTSS	1.45	4.47
25	402808			ENSP00000235229:SEMB.	1.00	1.87
	403021			C21000030:gi 9955960 ref NP_063957.1 AT	1.00	149.00
	403438			NM_031419*Homo sapiens molecule possess	1.06	2.96
	403687			NM_007037*Homo sapiens a disintegrin-li	0.04	4.89
30	403764			NM_005463:Homo sapiens heterogeneous nuc	1.00	225.00
	404277			NM_019111*Homo sapiens major histocompa	0.97	1.93
	404288			NM_002944*Homo sapiens v-ros avian UR2	1.00	68.00
	404518	AI815601		CD83 antigen (activated B lymphocytes, i	0.02	1.83
	405106			C11001637*gi 5032241 ref NP_005732.1 z	1.00	235.00
35	405381			Target Exon	1.00	93.00
	406387			Target Exon	1.37	6.02
	406646	M33600		major histocompatibility complex, class	0.88	2.46
	406714	AI219304	Hs.266959	hemoglobin, gamma G	0.01	3.19
	406753	AA505665	Hs.217493	annexin A2	1.00	147.00
40	406973	M34996	Hs.198253	major histocompatibility complex, class	1.03	2.04
	407248	U82275	Hs.94498	leukocyte immunoglobulin-like receptor,	1.00	64.00
	407510	U96191		gb:Human trophoblast hypoxia-regulated f	1.00	90.00
	407731	NM_000066	Hs.38069	complement component 8, beta polypeptide	1.00	67.00
	407830	NM_001086	Hs.587	arylacetamide deacetylase (esterase)	1.00	102.00
45	408045	AW138959	Hs.245123	ESTs	1.00	70.00
	408074	R20723		ESTs	1.00	112.00
	408374	AW025430	Hs.155591	forkhead box F1	0.07	10.17
	409064	AA062954	Hs.141883	ESTs	0.39	2.31
	409083	AF050083	Hs.673	interleukin 12A (natural killer cell sti	1.00	95.00
50	409153	W03754	Hs.50813	hypothetical protein FLJ20022	0.01	4.55
	409203	AA780473	Hs.687	cytochrome P450, subfamily IVB, polypept	0.01	3.72
	409238	AL049990	Hs.51515	Homo sapiens mRNA; cDNA DKFZp564G112 (fr	1.00	79.00
	409389	AB007979	Hs.301281	Homo sapiens mRNA, chromosome 1 specific	0.14	27.35
	409718	D86640	Hs.56045	src homology three (SH3) and cysteine ri	1.00	113.00
55	410798	BE178622	Hs.16291	gb:PM3-HT0605-270200-001-a02 HT0605 Homo	0.64	2.47
	411020	NM_006770	Hs.67726	macrophage receptor with collagenous str	0.55	2.40
	411667	BE160198		gb:QV1-HT0413-010200-059-h03 HT0413 Homo	1.00	111.00
	412000	AW576555	Hs.15780	ATP-binding cassette, sub-family A (ABC1	1.00	95.00
	412358	BE047490	Hs.24172	ESTs	1.00	87.00
60	412420	AL035668	Hs.73853	bone morphogenetic protein 2	1.43	8.07
	412564	X83703	Hs.31432	cardiac ankyrin repeat protein	0.02	3.07
	412869	AA290712	Hs.82407	CXC chemokine ligand 16	0.93	1.72
	412870	N22788	Hs.82407	CXC chemokine ligand 16	0.97	1.51
	413529	U11874	Hs.846	interleukin 8 receptor, beta	0.02	2.42
65	413533	BE146973		gb:QV4-HT0222-011199-019-e05 HT0222 Homo	0.65	1.50
	413689	BE157286	Hs.20631	zinc finger protein, subfamily 1A, 5 (Pe	20.87	232.00
	413724	AA131466	Hs.23767	hypothetical protein FLJ12666	1.00	80.00
	413800	AI129238	Hs.192235	ESTs	1.00	85.00
	413802	AW964490	Hs.32241	ESTs, Weakly similar to S65657 alpha-1C-	1.00	213.00
70	413829	NM_001872	Hs.75572	carboxypeptidase B2 (plasma)	0.02	3.93
	414376	BE393856	Hs.66915	ESTs, Weakly similar to 16.7Kd protein [1.00	115.00
	414577	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to	0.49	1.94
	414700	H63202	Hs.38163	ESTs	0.03	3.75
	415078	AA311223	Hs.283091	found in inflammatory zone 3	0.86	1.95
75	415120	N64464	Hs.34950	ESTs	1.00	120.00
	415323	BE269352	Hs.949	neutrophil cytosolic factor 2 (65kD, chr	0.60	2.48
	415335	AA847758	Hs.111030	ESTs	1.00	95.00
	415582	W92445	Hs.165195	Homo sapiens cDNA FLJ14237 fis, clone NT	1.00	136.00
	416030	H15251	Hs.21948	ESTs	0.02	8.07
80	416427	BE244050	Hs.79307	Rac/Cdc42 guanine exchange factor (GEF)	1.00	73.00
	416464	NM_000132	Hs.79345	coagulation factor VIII, procoagulant co	0.70	3.35
	416585	X54162	Hs.79386	leiomodulin 1 (smooth muscle)	0.06	6.56
	416847	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do	0.70	3.66
	417148	AA359896	Hs.283885	hypothetical protein FLJ14902	1.00	114.00
85	417370	T28651	Hs.82030	tryptophanyl-tRNA synthetase	0.85	1.30
	417673	T87281	Hs.16355	ESTs	0.15	15.54

	418067	AI127958	Hs.83393	cystatin E/M	0.81	1.74
	418295	C01566	Hs.85671	ESTs	1.00	99.00
	418543	J03798	Hs.86948	small nuclear ribonucleoprotein D1 polyp	1.00	60.00
	418832	X04011	Hs.88974	cytochrome b-245, beta polypeptide (chro	2.40	14.74
5	418945	BE246762	Hs.89499	arachidonate 5-lipoxygenase	0.67	3.16
	419261	X07876	Hs.89791	wingless-type MMTV integration site fami	1.00	73.00
	419564	U08989	Hs.91139	solute carrier family 1 (neuronal/epith	1.00	192.00
	419574	AK001989	Hs.91165	hypothetical protein	1.00	94.00
10	419968	X04430	Hs.93913	interleukin 6 (interferon, beta 2)	61.16	500.00
	420256	U84722	Hs.76205	cadherin 5, type 2, VE-cadherin (vascula	0.52	1.70
	420285	AA258124	Hs.293878	ESTs, Moderately similar to ZN91_HUMAN Z	1.00	172.00
	420577	AA278436	Hs.186649	ESTs	1.00	97.00
	421262	AA286746	Hs.9343	Homo sapiens cDNA FLJ14265 fis, clone PL	1.00	64.00
15	421445	AA913059	Hs.104433	Homo sapiens, clone IMAGE:4054868, mRNA	0.88	1.51
	421470	R27496	Hs.1378	annexin A3	0.05	11.26
	421478	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos	1.00	73.00
	421563	NM_006433	Hs.105806	granulysin	0.82	2.42
	421566	NM_000399	Hs.1395	early growth response 2 (Krox-20 (Drosop	5.50	31.57
20	421855	F06504	Hs.27384	ESTs, Moderately similar to ALU4_HUMAN A	1.00	129.00
	421913	AI934365	Hs.109439	osteoglycin (osteoinductive factor, mime	1.00	101.00
	421952	AA300900	Hs.98849	ESTs, Moderately similar to AF161511 1 H	0.60	63.60
	422232	D43945	Hs.113274	transcription factor EC	1.00	148.00
	422386	AF105374	Hs.115830	heparan sulfate (glucosamine) 3-O-sulfot	1.40	3.98
25	423168	R34385	Hs.124940	GTP-binding protein	0.34	3.59
	423196	AK001866	Hs.125139	hypothetical protein FLJ11004	0.55	2.00
	423387	AJ012074		vasoactive intestinal peptide receptor 1	0.09	2.13
	423424	AF150241	Hs.128433	prostaglandin D2 synthase, hematopoietic	1.00	141.00
	423456	AL110151	Hs.128797	DKFZP586D0824 protein	1.00	66.00
30	423695	Z92546		Sushi domain (SCR repeat) containing	0.73	1.27
	424027	AW337575	Hs.201591	ESTs	0.54	2.58
	424212	NM_005814	Hs.143131	glycoprotein A33 (transmembrane)	0.77	2.47
	425087	R62424	Hs.126059	ESTs	1.00	74.00
	425175	AF020202	Hs.155001	UNC13 (C. elegans)-like	0.85	1.96
35	425771	BE561776	Hs.159494	Bruton agammaglobulinemia tyrosine kinas	1.18	2.56
	426486	BE178285	Hs.170056	Homo sapiens mRNA; cDNA DKFZP586B0220 (f	1.00	76.00
	427507	AF240467	Hs.179152	tol-like receptor 7	1.00	63.00
	427618	NM_000760	Hs.2175	colony stimulating factor 3 receptor (gr	0.60	2.19
	427732	NM_002980	Hs.2199	secretin receptor	0.97	1.42
40	427952	AA765368	Hs.293941	ESTs, Moderately similar to A53959 throm	1.00	105.00
	428709	BE268717	Hs.104916	hypothetical protein FLJ21940	1.00	80.00
	428769	AW207175	Hs.106771	ESTs	0.09	2.55
	428780	AJ478578	Hs.50636	ESTs	1.00	98.00
	428833	AJ928355	Hs.185805	ESTs	1.00	113.00
45	429657	D13626	Hs.2465	KIAA0001 gene product; putative G-protei	1.00	52.00
	430212	AA469153		gb:nc67f04.s1 NC1_CGAP_P1 Homo sapiens	1.00	132.00
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	0.11	15.60
	430376	AW292053	Hs.12532	chromosome 1 open reading frame 21	1.00	103.00
50	430414	AW365665	Hs.120388	ESTs	0.50	6.96
	430656	AA482900	Hs.162080	ESTs	1.00	70.00
	430843	AJ734149	Hs.119514	ESTs	1.00	90.00
	430998	AF128847	Hs.204038	indolethylamine N-methyltransferase	0.29	1.84
	431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	1.00	79.00
	431921	N46466	Hs.58879	ESTs	0.91	1.67
55	432176	AW090386	Hs.112278	arrestin, beta 1	0.65	2.63
	432203	AA305746	Hs.49	macrophage scavenger receptor 1	1.00	76.00
	432231	AA339977	Hs.274127	CLST 11240 protein	0.46	1.46
	432485	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)	0.79	2.25
	432522	D11466	Hs.51	phosphatidylinositol glycan, class A (pa	1.93	4.83
60	432596	AJ224741	Hs.278461	matrilin 3	0.04	5.79
	432850	X87723	Hs.3110	angiotensin receptor 2	1.00	167.00
	433138	AB029496	Hs.59729	semaphorin sem2	0.04	9.16
	433563	AJ732637	Hs.277901	ESTs	1.00	91.00
	433588	AI056872	Hs.133386	ESTs	120.16	315.00
65	434445	AJ349306	Hs.11782	ESTs	0.60	1.84
	435496	AW840171	Hs.265398	ESTs, Weakly similar to transformation-r	1.00	128.00
	435974	U29690	Hs.37744	Homo sapiens beta-1 adrenergic receptor	1.00	108.00
	436061	AJ248584	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone C	1.00	91.00
	437157	BE048860	Hs.120655	ESTs	1.00	87.00
	437207	T27503	Hs.15929	hypothetical protein FLJ12910	1.00	105.00
70	437311	AA370041	Hs.9456	SWI/SNF related, matrix associated, acti	1.00	71.00
	437439	H29796	Hs.269622	ESTs	1.00	115.00
	438199	AW016531	Hs.122147	ESTs	1.00	80.00
	439551	W72062	Hs.111112	ESTs	0.30	3.10
75	440515	AJ131245	Hs.7239	SEC24 (S. cerevisiae) related gene fami	1.00	77.00
	440687	AJ799488	Hs.135905	ESTs	1.00	85.00
	441025	AA913880	Hs.176379	ESTs	1.00	82.00
	441384	AA447849	Hs.288660	Homo sapiens cDNA: FLJ22182 fis, clone H	0.79	1.89
	441735	AJ738675	Hs.127346	ESTs	1.00	75.00
80	442200	AW590572	Hs.235768	ESTs	0.78	5.83
	442832	AW206560	Hs.253569	ESTs	0.03	10.88
	442957	AJ949952	Hs.49397	ESTs	1.00	70.00
	443282	T47764	Hs.132917	ESTs	1.00	197.00
	443547	AW271273	Hs.23767	hypothetical protein FLJ12666	1.00	253.00
	443951	F13272	Hs.111334	ferritin, light polypeptide	0.55	2.09
85	444330	AI597655	Hs.49265	ESTs	1.00	90.00

	444515	AW204308	Hs.169979	ESTs	1.00	84.00
	445769	AI741471	Hs.23666	ESTs	0.02	4.38
	445908	R13580	Hs.13436	Homo sapiens clone 24425 mRNA sequence	1.00	97.00
	446291	BE397753	Hs.14623	Interferon, gamma-inducible protein 30	0.93	1.69
5	446917	AI347863	Hs.156572	ESTs	1.00	106.00
	447261	NM_006691	Hs.17917	extracellular link domain-containing 1	0.40	47.20
	447432	AW958473	Hs.301957	nudix (nucleoside diphosphate linked mol	1.00	100.00
	447482	AB033059	Hs.18705	KIAA1233 protein	0.05	8.21
	447997	H00656	Hs.29792	ESTs, Weakly similar to I38022 hypotheti	0.02	5.42
10	448299	AA497044	Hs.20887	hypothetical protein FLJ10392	1.00	79.00
	448782	AL050295	Hs.22039	KIAA0758 protein	0.42	1.56
	450575	NM_005859	Hs.29117	purine-rich element binding protein A	0.17	11.33
	450584	AA040403	Hs.60371	ESTs	1.00	94.00
	450693	AW450461	Hs.203965	ESTs	1.00	91.00
15	450715	AI266484	Hs.31570	ESTs, Weakly similar to KIAA1324 protein	1.00	152.00
	451103	R52804	Hs.25956	DKFZP564D206 protein	1.00	86.00
	451220	AF124251	Hs.26054	novel SH2-containing protein 3	0.60	1.30
	451668	Z43948	Hs.326444	cartilage acidic protein 1	0.54	1.91
	452197	AW023595	Hs.232048	ESTs	1.00	67.00
20	452331	AA598509	Hs.29117	purine-rich element binding protein A	4.53	11.07
	452353	C18825	Hs.29191	epithelial membrane protein 2	0.72	2.24
	453049	BE537217	Hs.30343	ESTs	1.00	68.00
	453107	NM_016113	Hs.279746	vanilloid receptor-like protein 1	0.83	1.70
	453355	AW295374	Hs.31412	Homo sapiens cDNA FLJ11422 fis, clone HE	1.00	132.00
25	453390	AA862496	Hs.28482	ESTs	1.00	72.00
	453531	AA417940		ESTs, Weakly similar to JC5795 CDEP prot	1.00	68.00
	454741	BE154396		gb:CM2-HT0342-091299-050-b05 HT0342 Homo	0.57	2.89
	456579	AA287827	Hs.284205	up-regulated by BCG-CWS	1.00	82.00
	456672	AK002016	Hs.114727	Homo sapiens, clone MGC:16327, mRNA, com	0.79	1.96
30	457400	AF032906	Hs.252549	cathepsin Z	1.03	3.25
	457718	F18572	Hs.22978	ESTs, Weakly similar to ALU4_HUMAN ALU S	1.00	113.00
	459696	F03027		gb:HSC1KA072 normalized Infant brain cDN	1.00	544.00

TABLE 10B

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

40	Pkey	CAT Number	Accession
	408074	103584_1	R20723 AA263003 AA333976 AA334725 AA334151 AW965490 AA310513 AJ810530 D31302 AW134897 AA830127 AA046953 AJ668930 C06094 AW104534
	411667	1253334_1	BE160198 AW935898 T11520 AW935930 AW856073 AW861034
45	413533	1375344_1	BE146973 BE146972 BE147042 BE147018 BE146783 BE147020 BE146781 BE147019 BE146766 BE147021 BE146952 BE146767 BE147044 BE146797 BE146776 BE146985 BE146793 BE146768 BE146771 BE146954 BE146760 BE147048 BE147025 BE147030
	423387	22779_1	AJ012074 U11087 L13288 X75299 L20295 AW630780 H14880 T28037 AI872991 R72136 AW449839 T81622 T79697 T29519 R94105 T83923 R73300 AI797007 R73390 AA961010 H74168 AI689932 BE045543 AI808418 AI608912 AI806573 AW884084 AW872978 AW872985 AA565655 AI022915 R50647 R73210 H45098 R46451 AW166269 T71132 AI254547 R52146 AI304920 R73391 AW884059 AW884085 H73241 T60038 T79612 R73145 R50549 AI094557 AI668793 R72302 AI564366 W01956 AA418962 W32571 R72840 H45409 R72085 R46356 R46758 AA508805 AA418798 T83751 R94072 T16182 AA928785 AA903896
50	423696	23112_1	Z92546 AA330586 AI570568 AW341487 AI827050 AW298668 AI792189 AI015693 AI733599 AI572251 AI672488 AW193262 AI244716 AI864375 AI206100 AA912444 AI269365 AI640254 AW772466 AI867336 AA627604 H16914 AA358477 AA338009
	430212	314437_1	AA469153 AI718503 AA469225
	436532	421802_1	AA721522 AW975443 T93070
55	453531	97026_1	AA417940 AA036735 T07025
	454741	1232559_1	BE154396 AW817959 BE154393

TABLE 10C

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA
 sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

	Pkey	Ref	Strand	Nt_position
	400754	7331445	Plus	144559-144684
	401045	8117619	Plus	90044-90184,91111-91345
70	401083	3242744	Plus	33192-33360
	402474	7547175	Minus	53526-53628,55755-55920,57530-57757
	402808	6456148	Minus	114964-115136,115461-115585,115931-116047,117666-117771,118004-118102
	403021	7547270	Plus	120799-120966
	403421	9665041	Minus	126609-126773,139986-140205
75	403438	9719679	Plus	90792-90938
	403687	7387384	Plus	9009-9534
	403764	7717105	Minus	118692-118853
	404277	1834458	Minus	91665-91946
	404288	2769644	Plus	3512-3691
80	404394	3135305	Minus	37121-37205,37491-37762,41053-41140,41322-41593,41773-41919
	404518	8151988	Plus	84494-84603
	404916	7341826	Plus	91057-91188
	405106	8079395	Minus	80877-81418
	405257	7329310	Plus	73121-73273
85	405381	6006920	Minus	7638-8054

WO 02/086443
406387 9256180 Plus

116229-116371,117512-117651

PCT/US02/12476

TABLE 11A: Genes Distinguishing Adenocarcinoma from Other Lung Diseases and Normal Lung

Table 11A shows about 84 genes upregulated in lung adenocarcinomas relative to other lung tumors, non-malignant lung disease, and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 11B show the accession numbers for those Pkey's lacking UnigeneID's for table 11A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 11C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 11A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigeneID: Unigene number
 Unigene Title: Unigene gene title
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
403329			Target Exon	1.00	61.00
406399			NM_003122*:Homo sapiens serine protease	1.00	39.00
406690	M29540	Hs.220529	cardiomyoblastic antigen-related cell ad	226.37	350.00
407869	A827376	Hs.24391	hypothetical protein FLJ13612	0.77	1.18
407881	AW072003	Hs.40968	heparan sulfate (glucosamine) 3-O-sulfat	1.00	10.00
408908	BE296227	Hs.250822	serine/threonine kinase 15	7.76	1.00
409103	AF251237	Hs.112208	XAGE-1 protein	80.44	40.00
409187	AF154830	Hs.50966	carbamoyl-phosphate synthetase 1, mitoch	1.00	1.00
409269	AA576953	Hs.22972	hypothetical protein FLJ13352	1.00	1.00
410076	T05387	Hs.7991	ESTs	1.12	1.50
410102	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	9.89	1.00
410399	BE068889		synuclein, gamma (breast cancer-specific	0.92	1.08
411908	L27943	Hs.72924	cytidine deaminase	1.00	1.00
412612	NM_000047	Hs.74131	arylsulfatase E (chondrodysplasia puncta	1.02	1.03
414075	U11862	Hs.75741	amiloride binding protein 1 (amine oxida	0.84	1.07
416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	3.67	1.00
417542	J04129	Hs.82269	progesterone-associated endometrial prote	1.28	1.35
419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	1.00	1.00
419502	AJ076704		fibrinogen, A alpha polypeptide	13.05	115.00
419631	AW188117	Hs.303154	popeye protein 3	1.00	13.00
420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	1.00	8.00
421155	H87879	Hs.102267	lysyl oxidase	1.00	15.00
421190	U95031	Hs.102482	mucin 5, subtype B, tracheobronchial	1.17	1.55
421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	1.46	1.76
421515	Y11339	Hs.105352	GalNAc alpha-2, 6-sialyltransferase 1, l	1.00	3.00
421582	A910275		trefol factor 1 (breast cancer, estrogen	1.23	1.00
422026	U80738	Hs.110826	trinucleotide repeat containing 9	1.00	52.00
422095	A868872	Hs.282804	hypothetical protein FLJ22704	4.37	2.34
422311	AF073515	Hs.114948	cytokine receptor-like factor 1	1.15	1.78
422867	L32137	Hs.1684	cartilage oligomeric matrix protein (pse	1.69	3.17
423472	AF041260	Hs.129057	breast carcinoma amplified sequence 1	48.13	72.00
423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	1.00	50.00
424502	AF242388	Hs.149585	lengsin	1.00	1.00
424544	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	1.00	59.00
424905	NM_002497	Hs.153704	NIMA (never in mitosis gene a)-related k	21.35	1.00
424960	BE245380	Hs.153952	5' nucleotidase (CD73)	1.00	1.00
425523	AB007948	Hs.158244	KIAA0479 protein	1.00	35.00
426230	AA367019	Hs.241395	protease, serine, 1 (trypsin 1)	1.00	83.00
427701	AA411101	Hs.243886	nuclear autoantigenic sperm protein (his	7.41	34.00
428585	AB007863	Hs.185140	KIAA0403 protein	1.00	6.00
428758	AA433988	Hs.98502	hypothetical protein FLJ14303	1.06	1.13
429170	NM_001394	Hs.2359	dual specificity phosphatase 4	16.18	105.00
429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	1.07	1.00
429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	1.59	1.69
430508	AJ015435	Hs.104637	ESTs	4.75	7.27
430985	AA490232	Hs.27323	ESTs, Weakly similar to I78885 serine/th	0.94	1.28
431548	AJ834273	Hs.9711	novel protein	5.66	15.00
431566	AF176012	Hs.260720	J domain containing protein 1	49.76	37.00
431986	AA536130	Hs.149018	Novel human gene mapping to chromosome 20	1.19	1.47
432375	BE536059	Hs.2962	S100 calcium-binding protein P	1.65	1.06
432677	NM_004482	Hs.278611	UDP-N-acetyl-alpha-D-galactosamine:polyp	1.00	48.00
433556	W56321	Hs.111460	calcium/calmodulin-dependent protein kin	1.00	19.00
433819	AW511097	Hs.112765	ESTs	3.71	8.00
434001	AW509005	Hs.3697	serine (or cysteine) proteinase inhibito	29.31	72.00
434424	AJ811202	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	1.00	64.00
434792	AA649253	Hs.132458	ESTs	8.52	44.00
436217	T53925	Hs.107	fibrinogen-like 1	57.97	31.00
436749	AA584890	Hs.5302	lectin, galactoside-binding, soluble, 4	1.10	1.41
436972	AA284679	Hs.25640	claudin 3	1.59	1.46
437866	AA156781		metallothionein 1E (functional)	3.62	101.00
437935	AW393591	Hs.5940	mucin 13, epithelial transmembrane	1.60	1.39
438915	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regi	1.00	1.00
439451	AF086270	Hs.278554	heterochromatin-like protein 1	23.28	52.00

	439759	AL359055	Hs.57709	Homo sapiens mRNA full length insert cDN	1.00	21.00
	441031	AI110684	Hs.7645	fibrinogen, B beta polypeptide	1.41	99.00
	441377	BE218239	Hs.202656	ESTs	22.03	1.00
5	443614	AV655386	Hs.7645	fibrinogen, B beta polypeptide	1.00	16.00
	443813	AA876372	Hs.93961	Homo sapiens mRNA: cDNA DKFZp667D095 (fr	1.20	1.99
	443991	NM_002250	Hs.10082	potassium intermediate/small conductance	5.71	6.87
	444670	H58373	Hs.332938	hypothetical protein MGC5370	1.98	38.00
	444931	AV652066	Hs.75113	general transcription factor IIIA	1.00	54.00
10	445102	AW168067	Hs.317694	ESTs	1.00	1.00
	446163	AA026880	Hs.25252	Homo sapiens cDNA FLJ13603 fis, clone PL	1.00	36.00
	446469	BE094848	Hs.15113	homogenisase 1,2-dioxygenase (homogeni	1.00	11.00
	447388	AW630534	Hs.76277	Homo sapiens, clone MGC:9381, mRNA, comp	1.24	1.16
	447532	AK000614	Hs.18791	hypothetical protein FLJ20607	1.23	1.63
	448243	AW369771	Hs.52620	integrin, beta 8	15.84	1.00
15	448844	AI581519	Hs.177164	ESTs	1.00	31.00
	449444	AW818436	Hs.23590	solute carrier family 16 (monocarboxylic	1.00	83.00
	451807	W52854		hypothetical protein FLJ23293 similar to	1.55	35.00
	452689	F33868	Hs.284176	transferrin	1.54	1.44
	453392	U23762	Hs.32964	SRY (sex determining region Y)-box 11	1.00	16.00
20	453464	AI884911	Hs.32989	receptor (calcitonin) activity modifying	1.55	2.45
	453735	AI066629	Hs.125073	ESTs	1.01	1.30

TABLE 11B

25 Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

30	Pkey	CAT Number	Accession
	410399	11995_1	BE068889 BE068882 AF044311 AF017256 NM_003087 AF037207 AF010126 AA633976 AA872836 BE298825 BE299889 AI016464 AI684600 AI936527 AA804675 AA394097 AI139533 AA946606 BE171313 AA722407 AA293803 AI458480 AA055035 AA055968 AW796957 AI637713 AA410737 H49348 AA485472 AA411094 AA235594 AA402624 AA443638 AW452137 AA421708 AW265211 AF493266 AA365132 AW966044 AU076704 T74854 T74860 T72068 T73265 T73873 T69180 T74658 T58786 T60385 T73410 T68781 T67845 T67593 T73952 T67864 T60630 T68367 T68401 T53959 T72360 T72099 T60377 T58961 T71712 T72821 T64738 T74645 T72037 T68668 T72063 T73258 T72826 T64242 T68220 T74673 T71800 T68355 T61227 T62738 T69317 T53850 T64692 T73768 T73962 T73382 T68914 T70975 T73400 T60631 T73277 T73203 T70498 T61409 T58925 NM_000508 M64982 T68301 T73729 T69445 T60424 T67922 T67736 T68716 T67755 T74765 T73819 T58719 T74766 T60477 T74863 T61109 T68329 T58850 T71857 T73425 T53736 T68607 T58898 T64309 T72031 T72079 T64305 T71909 T68107 T71916 T73787 T56035 T64425 T71870 T60476 T61376 T67820 T71895 T41006 T69441 T68170 T74617 T71958 T69440 T61875 T60796 T48353 T71914 T53939 T64121 AA693996 T72525 T67779 T68078 AA011465 AA345378 AV654847 AV654272 AV656001 AI064740 T82697 N33594 AA344542 AW805054 AI207457 T61743 AA026737 H94389 AA382695 AA918409 T68044 S82092 T39959 AI017721 AA312395 AA312919 T40156 H66239 AV652989 H38728 R98521 AV656200 R95790 W03250 W00913 AA344136 AV660126 R97923 AA343596 AW470774 AV651256 N54417 AA812862 AW182929 AI111192 H61463 H72060 AA344503 H38639 AI277511 AV661108 AI207625 T47810 AA235252 T27853 T47778 R95746 H70620 AA701463 AW827166 R98475 C20925 AV657287 T71959 T71313 T73920 T73333 T61618 T69293 T69283 T73931 T72178 T72456 AV645639 AV653476 T72957 T72300 T58905 T71457 T70494 T72956 T70495 T68267 T74407 T85778 AA344726 T27854 T74485 T74101 T73868 T71518 T72304 AA343953 T73909 T68070 T72065 H72149 T73493 T73495 AV645993 R02293 T70475 T64751 AA344441 AA343657 AA345732 AA344328 AI110639 AA344603 AF063513 T64696 T68516 T72223 T60507 T67633 R29500 T72517 R02292 T60599 T69206 T70452 T74677 R29366 T61277 T74914 T60352 R29675 T74843 AV645792 AA344408 T69197 T72057 T69368 T69359 T68258 AV650429 T73341 T61702 T74598 T40095 K02272 T40106 AA343045 AA341908 AA341907 AA342807 AA341964 T53747 T72042 T62764 AI054899 AA343060 T67832 T72440 T71770 T68091 T69108 T72449 T69167 T71289 T68251 AV654844 T64375 AA345234 T67598 AA011414 T68036 H48262 AI207557 T68219 W86031 T69081 T64232 R93196 T62136 AV650539 H67459 T72978 AA344583 T60362 H58121 T95711 T72803 T68055 T71715 R29036 T72793 T69122 T64595 T62888 T69139 T68291 T64652 T67971 T46862 AA693592 AI248502 R29454 T64764 T57001 T73052 T71429 T51176 T58856 AV655414 H90426 AA342489 T73666 T67848 T72512 T53835 T67837 T73317 T74273 T69420 T68245 T74380 T67862 T74474 T56068 AI910275 X00474 X52003 X05030 NM_003225 AA314326 AA309400 AA506787 AA314825 AI571948 AA507595 AA614579 AA587613 R83818 AA588312 AA614409 AA307578 AI925552 AW950155 AI910083 M12075 BE074052 AW004668 AA578574 AA582084 BE074053 BE074126 BE074140 AA514776 AA588034 BE074051 BE074068 AW009769 AW050690 AA858276 R55389 AI001051 AW050700 AW750216 AA614539 BE074045 AI307407 AW602303 BE073575 AI202532 AA524242 AI970839 AI909751 BE076078 AI909749 R55292 AA156781 AW293839 U52054 AA024963 AA778446 BE073977 AW444904 AW602574 BE164040 BE164012 BE163972 BE163974 BE163992 AA837481 AW468444 BE185091 AW468002 AA687333 AA811830 AA581806 AI666686 AI572124 AA043777 AA040926 D02160 AI536733 AA812489 AW874142 AI471883 W84421 AA156850 W52854 AL117600 BE208116 BE208432 BE206239 BE082291 AW953423 AA351619 BE180648 BE140560 W60080 AA865478 N90291 AW450652 AW449519 AA993634 AI906539 AA351618 AW449522 AI827626 AA904788 AA380381 AA896045 AA774409 BE003229 Z41756

TABLE 11C

70 Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

75	Pkey	Ref	Strand	Nt_position
	403329	8516120	Plus	96450-96598
	406399	9256288	Minus	63448-63554

TABLE 12A: Genes Distinguishing Squamous Cell Carcinoma from Other Lung Diseases and Normal Lung

Table 12A shows about 72 genes upregulated in squamous cell carcinomas of the lung relative to other lung tumors, non-malignant lung disease, and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu33 Genechip array.

Table 12B show the accession numbers for those Pkey's lacking UnigenelD's for table 12A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 12C show the genomic positioning for those Pkey's lacking UnigenelD's and accession numbers in table 12A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: UnigenelD number
 Unigenel Title: Unigenel gene title
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigenelD	Unigenel Title	R1	R2
400289	X07820	Hs.2259	matrix metalloproteinase 10 (stromelysin	132.45	4.00
400666			NM_002425:Homo sapiens matrix metallopro	3.26	3.22
401780			NM_005557:Homo sapiens keratin 16 (foca	26.47	10.50
401781			Target Exon	10.33	4.61
401785			NM_002275:Homo sapiens keratin 15 (KRT1	4.13	2.70
401994			Target Exon	61.84	47.00
402075			ENSP00000251056:Plasma membrane calcium	1.00	1.00
404996			Target Exon	1.00	1.00
407839	AA045144	Hs.161566	ESTs	173.91	108.00
408000	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	151.17	8.00
408522	AI541214	Hs.46320	Small proline-rich protein SPRK (human,	1.98	1.24
410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	10.04	1.00
415091	AL044872	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A sy	1.00	30.00
415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	24.30	1.00
416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	53.29	51.00
417034	NM_006183	Hs.80962	neurotensin	1.00	1.00
417366	BE185289	Hs.1076	small proline-rich protein 1B (cornelin)	8.97	3.27
418663	AK001100	Hs.41690	desmocollin 3	112.17	19.00
418678	NM_001327	Hs.87225	cancer/testis antigen	1.18	1.10
419121	AA374372	Hs.89626	parathyroid hormone-like hormone	1.00	1.00
420783	AI659338	Hs.99923	lectin, galactoside-binding, soluble, 7	3.04	1.25
421773	W69233	Hs.112457	ESTs	1.12	1.14
421948	L42583	Hs.334309	keratin 6A	51.83	20.25
421978	AJ243662	Hs.110196	NICE-1 protein	1.01	0.91
422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	2.37	1.10
422440	NM_004812	Hs.116724	aldo-keto reductase family 1, member B10	47.53	32.00
423634	AW59908	Hs.1690	heparin-binding growth factor binding pr	76.02	1.00
423725	AJ403108	Hs.132127	hypothetical protein LOC57822	4.20	1.00
423738	AB002134	Hs.132195	airway trypsin-like protease	10.14	51.00
424012	AW368377	Hs.137559	tumor protein 63 kDa with strong homolog	233.42	68.00
424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibito	1.00	1.00
424098	AF077374	Hs.139322	small proline-rich protein 3	137.82	54.00
424834	AK001432	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	56.19	12.00
425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	33.45	1.00
427099	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	4.24	17.00
427335	AA448542	Hs.251677	G antigen 7B	51.83	4.00
428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN G ANT	1.00	1.00
428645	AA431400	Hs.98729	ESTs, Weakly similar to 2017205A dihydro	1.00	16.00
428748	AW593206	Hs.98785	Ksp37 protein	1.00	87.00
429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	2.01	1.18
429538	BE182592	Hs.11261	small proline-rich protein 2A	4.43	2.90
429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	11.80	1.00
430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	12.28	41.00
430890	X54232	Hs.2699	glypican 1	1.58	1.40
431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	60.25	28.00
431846	BE019924	Hs.271580	uroplakin 1B	4.49	2.51
433091	Y12642	Hs.3185	lymphocyte antigen 6 complex, locus D	1.20	1.09
434360	AW015415	Hs.127780	ESTs	40.98	27.00
434880	U02388	Hs.101	cytochrome P450, subfamily IVF, polypept	1.00	1.00
435505	AF200492	Hs.211238	interleukin-1 homolog 1	1.00	38.00
435793	AB037734	Hs.4993	KIAA1313 protein	23.68	42.00
436511	AA721252	Hs.291502	ESTs	16.76	14.00
438403	AA806607	Hs.292206	ESTs	1.00	1.00
439285	AL133916		hypothetical protein FLJ20093	46.23	139.00
439606	W79123	Hs.58561	G protein-coupled receptor 87	33.61	1.00
439670	AF088076	Hs.59507	ESTs, Weakly similar to AC004858 3 U1 sm	1.00	1.00
439706	AW872527	Hs.59761	ESTs, Weakly similar to DAPI_HUMAN DEATH	86.55	11.00
440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma	62.88	147.00
441525	AW241867	Hs.127728	ESTs	1.53	1.42
443162	T49951	Hs.9029	DKFZP434G032 protein	31.11	38.00
444378	R41339	Hs.12569	ESTs	1.00	1.00

WO 02/086443

PCT/US02/12476

446292	AF031497	Hs.279682	Rh type C glycoprotein	1.55	1.25
447078	AW985727	Hs.9914	ESTs	47.24	24.00
447342	AI199258	Hs.19322	Homo sapiens, Similar to RUKEN cDNA 2010	28.63	1.00
449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	1.00	1.00
449101	AA205847	Hs.23016	G protein-coupled receptor	2.58	27.00
450832	AW970602	Hs.105421	ESTs	25.17	36.00
452240	AI591147	Hs.61232	ESTs	13.42	1.00
453317	NM_002277	Hs.41696	keratin, hair, acidic, 1	1.19	1.27
453830	AA534296	Hs.20953	ESTs	24.92	25.00
454098	W27953	Hs.29211	ESTs, Highly similar to S60712 band-6-pr	1.26	1.11
455501	AI358580	Hs.816	SRY (sex determining region Y)-box 2	206.11	1.00

TABLE 12B

15 Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

20 Pkey CAT Number Accession
 439285 47055_1 AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 AI346341 AI867454 N54784 AI655270 AI421279 AW014882
 AA775552 N62351 N59253 AA626243 AI341407 BE175639 AA456968 AI358918 AA457077

TABLE 12C

25 Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA
 sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 NL_position: Indicates nucleotide positions of predicted exons.

30 Pkey Ref Strand NL_position
 400666 8118496 Plus 17982-18115,20297-20456
 401780 7249190 Minus 28397-28617,28920-29045,29135-29296,29411-29557,29705-29787,30224-30573
 401781 7249190 Minus 83215-83435,83531-83658,83740-83901,84237-84393,84955-85037,86290-86814
 401785 7249190 Minus 165776-165996,166189-166314,166408-166569,167112-167268,167387-167469,168634-168942
 401994 4153858 Minus 42904-43124,43211-43336,44607-44763,45199-45281,46337-46732
 402075 8117407 Plus 121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
 404996 6007890 Plus 37999-38145,38652-38998,39727-39872,40557-40674,42351-42450

TABLE 13A: Genes Distinguishing Non-Malignant Lung Disease from Lung Tumors and Normal lung

Table 13A shows about 23 genes upregulated in non-malignant lung disease relative to lung tumors and normal lung. These genes were selected from about 59680 probesets on the Eos/Alfymetrix HuQ3 Genechip array.

Table 13B show the accession numbers for those Pkey's lacking UnigeneID's for table 13A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 13C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 13A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigeneID: Unigene number
 Unigene Title: Unigene gene title
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
408562	AJ436323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,	1.00	230.00
409031	AA376836	Hs.76728	ESTs	1.00	128.00
412372	R65998	Hs.285243	hypothetical protein FLJ22029	1.00	173.00
415910	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	1.00	145.00
417511	AL049176	Hs.82223	chordin-like	1.00	179.00
418819	AA228776	Hs.191721	ESTs	1.00	140.00
422060	R20893	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	1.00	156.00
424585	AA464840	Hs.131987	ESTs	1.00	167.00
426753	T89832	Hs.170278	ESTs	1.00	141.00
429496	AA453800	Hs.192793	ESTs	1.00	138.00
430719	AA488988	Hs.293796	ESTs	1.00	133.00
431089	BE041395		ESTs, Weakly similar to unknown protein	23.32	941.00
431385	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A	1.00	157.00
431728	NM_007351	Hs.268107	multimerin	1.00	157.00
436532	AA721522		gbzv54h12.r1 NCI_OGAP_Ew1 Homo sapiens	1.00	218.00
437960	AI669586	Hs.222194	ESTs	1.00	147.00
438202	AW169287	Hs.22588	ESTs	1.00	141.00
441499	AW298235	Hs.101689	ESTs	1.00	167.00
444513	AL120214	Hs.7117	glutamate receptor, ionotropic, AMPA 1	1.00	151.00
448253	H25899	Hs.201591	ESTs	1.00	141.00
453636	R67837	Hs.169872	ESTs	1.00	116.00
458332	AI000341	Hs.220491	ESTs	1.00	192.00
459587	AA031956		gbzkl15e04.s1 Soares_pregnant_uterus_NbH	1.00	154.00

TABLE 13B

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT Number	Accession
431089	327825_1	BE041395 AA491826 AA621946 AA715980 AA666102
436532	421802_1	AA721522 AW975443 T93070

TABLE 13C

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham L. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham L. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 NL_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125572-126076

TABLE 14A: Preferred Utility and Subcellular Localization for Potential Lung Disease Targets

Table 14A shows the subcellular localization and preferred utility for the genes appearing in Tables 9A and 10A. mAb symbolizes monoclonal antibody, diag symbolizes diagnostic, s.m. symbolizes small molecule, and CTL symbolizes cytotoxic lymphocyte ligand. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 14B shows the accession numbers for those Pkey's lacking UnigenelD's for table 14A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the 'Accession' column.

Table 14C shows the genomic positioning for those Pkey's lacking UnigenelD's and accession numbers in table 14A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number

ExAccn: Exemplar Accession number, Genbank accession number

UnigenelD: Unigenel number

Unigenel Title: Unigenel gene title

Pref.Utility: Preferred Utility

Pred.Loc: Predicted subcellular localization

Pkey	ExAccn	UnigenelD	Unigenel Title	Pref Utility	Pred. Loc
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin)	mAb & diag & s.m.	extracellular
400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	mAb	plasma membrane
402075			ENSP00000251056:Plasma membrane calcium	mAb & diag	secreted
407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagonist	diag	secreted
408243	Y00787	Hs.624	interleukin 8	diag	secreted
408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	mAb & s.m.	plasma membrane
408908	BE296227	Hs.250822	serine/threonine kinase 15	s.m.	cytoplasm
409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	CTL & diag	secreted
409103	AF251237	Hs.112208	XAGE-1 protein	CTL	nuclear
409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini	diag	secreted
409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibitor	diag	secreted
409757	NM_001898	Hs.123114	cystatin SN	diag	extracellular
409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.	CTL	nuclear
409956	AW103364	Hs.727	Inhibin, beta A (activin A, activin AB a	diag	extracellular
410001	AB041036	Hs.57771	kallikrein 11	diag	extracellular
410407	X66839	Hs.63287	carbonic anhydrase IX	mAb & s.m.	plasma membrane
410418	D31382	Hs.63325	transmembrane protease, serine 4	mAb & diag & s.m.	plasma membrane
412140	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rakkin	s.m.	
412719	AW016610	Hs.816	ESTs	s.m.	nuclear
414774	X02419	Hs.77274	plasminogen activator, urokinase	diag	extracellular
414883	AA926960		CDC28 protein kinase 1	s.m.	
415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	CTL & diag	extracellular
415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor	mAb & diag & s.m.	secreted
415817	U89867	Hs.78867	protein tyrosine phosphatase, receptor-t	mAb & s.m.	plasma membrane
416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	diag	extracellular
417034	NM_006183	Hs.80962	neurotensin	diag	extracellular
417079	U65590	Hs.81134	Interleukin 1 receptor antagonist	diag	extracellular
417308	H50720	Hs.81892	KIAA0101 gene product	s.m.	mitochondrial
417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor	mAb & diag	secreted
417433	BE270266	Hs.82128	ST4 oncogene related trophoblast glycoprotein	mAb	plasma membrane
417933	X02308	Hs.82962	thymidylate synthetase	s.m.	endoplasmic reticulum
418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me	s.m.	cytoplasm
418506	AA084248	Hs.85339	G protein-coupled receptor 39	mAb & s.m.	plasma membrane
418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)	CTL	cytoplasmic
419121	AA374372	Hs.89626	parathyroid hormone-like hormone	diag	secreted
419171	NM_002846	Hs.89655	protein tyrosine phosphatase, receptor t	mAb & s.m.	plasma membrane
419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	CTL & s.m.	mitochondrial
419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy	diag	secreted
419235	AW470411	Hs.288433	neurotrophin	mAb & diag	plasma membrane
419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	mAb & s.m.	plasma membrane
419556	U29615	Hs.91093	chitinase 1 (chitotriosidase)	mAb & diag	extracellular
420610	AI683183	Hs.99348	distal-less homeo box 5	CTL	nuclear
421110	AJ250717	Hs.1355	cathepsin E	sm & diag	extracellular
421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy	diag	secreted
421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	mAb & s.m.	plasma membrane
421552	AF026692	Hs.105700	secreted frizzled-related protein 4	diag	secreted
421753	BE314828	Hs.107911	ATP-binding cassette, sub-family B (MDR/	mAb & s.m.	plasma membrane
421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	mAb & s.m.	plasma membrane
422109	S73265	Hs.1473	gastrin-releasing peptide	diag	secreted
422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	diag	secreted
422282	AF019225	Hs.114309	apolipoprotein L	diag	secreted
422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	s.m.	nuclear
422424	AI186431	Hs.296538	prostate differentiation factor	diag	extracellular
422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	s.m.	cytoplasm
422809	AK001379	Hs.121028	hypothetical protein FLJ10549	s.m.	nuclear
422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	diag	extracellular
422956	BE545072	Hs.122579	ECT2 protein (Epithelial cell transformi	CTL & s.m.	
423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	diag	
423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	mAb & diag & s.m.	secreted
423951	D13666	Hs.136348	perlestin (OSF-2os)	mAb & diag	extracellular
424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibitor	diag	secreted
424381	AA285249	Hs.146329	protein kinase Chk2	s.m.	nuclear

	424502	AF242388	Hs.149585	lengsin	s.m.	cytoplasmic
	424503	NM_002205	Hs.149609	integrin, alpha 5 (fibronectin receptor,	mAb & s.m.	plasma membrane
	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	diag	extracellular
5	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	mAb & diag & s.m.	secreted
	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	s.m.	cytoplasmic
	425650	NM_001944	Hs.1925	desmoglein 3 (permpigus vulgaris antigen	mAb	plasma membrane
	425734	AF056209	Hs.159395	peptidylglycine alpha-amidating monooxyg	s.m.	
	425776	U25128	Hs.159499	parathyroid hormone receptor 2	mAb & diag	plasma membrane
10	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member	mAb & s.m.	plasma membrane
	426215	AW963419	Hs.155223	stanniocalcin 2	mAb & diag	secreted
	426427	M86699	Hs.169840	TTK protein kinase	CTL & s.m.	nuclear
	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic	mAb & diag	secreted
	427335	AA448542	Hs.251677	G antigen 7B	CTL	cytoplasmic
	427747	AW411425	Hs.180655	serine/threonine kinase 12	s.m.	cytoplasmic
15	428242	H55709	Hs.2250	leukemia inhibitory factor (cholestergic	diag	
	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	mAb & diag & s.m.	extracellular
	428450	NM_014791	Hs.184339	KIAAD175 gene product	s.m.	nuclear
	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to	s.m.	nuclear
20	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino	mAb & s.m.	plasma membrane
	428664	AK001666	Hs.189095	similar to SALL1 (sal (Drosophila)-like	CTL & s.m.	nuclear
	428698	AA852773	Hs.334838	KIAA1866 protein	mAb	
	428748	AW593206	Hs.98785	Ksp37 protein	diag	extracellular
	428758	AA433988	Hs.98502	CA125 antigen; mucin 16	diag	mitochondria*
25	428969	AF120274	Hs.194689	artemin	diag	extracellular
	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	mAb & s.m.	plasma membrane
	429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	mAb & s.m.	plasma membrane
	429547	AW009166	Hs.99376	ESTs	diag	secreted
	429510	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	mAb & diag	secreted
30	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	s.m.	
	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	mAb & s.m.	plasma membrane
	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu	diag	extracellular
	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha	mAb & s.m.	plasma membrane
	431846	BE019924	Hs.271580	uropod 1B	mAb & diag	plasma membrane
	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	mAb & diag	plasma membrane
35	432201	AF538613	Hs.298241	Transmembrane protease, serine 3	mAb & diag & s.m.	plasma membrane
	433001	AF217513	Hs.279905	clone HQ0310 PRO0310p1	s.m.	nuclear
	435505	AF200492	Hs.211238	interleukin-1 homolog 1	diag	secreted
	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	s.m.	
40	437016	AU076916	Hs.5398	guanine monophosphate synthetase	s.m.	cytoplasm
	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an	CTL	ER
	437789	AU81344	Hs.127812	ESTs, Weakly similar to T17330 hypoteli	CTL	nuclear
	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365O12.1 [H.s	mAb & s.m.	plasma membrane
	439223	AW238299	Hs.250618	UL16 binding protein 2	mAb	plasma membrane
45	439477	W69813	Hs.58042	ESTs, Moderately similar to GFR3_HUMAN G	mAb & s.m.	
	439606	W79123	Hs.58561	G protein-coupled receptor 87	mAb & s.m.	plasma membrane
	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	mAb & s.m.	plasma membrane
	440006	AK000517	Hs.6844	NALP2 protein; PYRIN-Containing APAF1-l	s.m.	nuclear
	441362	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E. coli Re	s.m.	
	442117	AW664964	Hs.128899	ESTs; hypothetical protein for IMAGE:447	mAb & s.m.	plasma membrane
50	443247	BE614387	Hs.333893	c-Myc target JPO1	CTL	extracellular*
	443426	AF098158	Hs.9329	chromosome 20 open reading frame 1	CTL	
	443859	NM_013409	Hs.9914	folistatin	diag	extracellular
	444006	BE395085	Hs.10086	type I transmembrane protein Fn14	mAb	plasma membrane
55	444371	BE540274	Hs.239	forkhead box M1	s.m.	nuclear
	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypoteli	diag	secreted
	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote	mAb & diag	plasma membrane
	445537	AJ245671	Hs.12844	EGF-like-domain, multiple 6	mAb & diag	secreted
	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	diag	secreted
	446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	diag	extracellular
60	447033	AJ357412	Hs.157601	ESTs	CTL & diag	secreted
	447342	AJ199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	CTL	
	448243	AW369771	Hs.52620	integrin, beta 8	mAb & s.m.	plasma membrane
	448844	AJ581519	Hs.177164	ESTs	mAb & s.m.	
65	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc	mAb	plasma membrane
	449722	BE280074	Hs.23960	cyclin B1	s.m.	cytoplasm
	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitte	mAb & s.m.	plasma membrane
	450375	AA009647	Hs.288467	a disintegrin and metalloproteinase doma	mAb & diag & s.m.	plasma membrane
	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-	mAb & diag	plasma membrane
70	450983	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	diag	secreted
	451668	Z43948	Hs.326444	cartilage acidic protein 1	mAb & diag	plasma membrane
	452281	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fis, clone PL	diag	
	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro	diag	extracellular
	452747	BE153855	Hs.61460	Ig superfamily receptor UNIR	mAb	plasma membrane
	452838	U65011	Hs.30743	preferentially expressed antigen in mela	CTL	nuclear
75	453968	AA847843	Hs.62711	High mobility group (nonhistone chromoso	CTL & s.m.	nuclear
	457489	AI693815	Hs.127179	cryptic gene	diag	secreted

TABLE 14B

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey CAT Number Accession

414883 15024_1 AA926960 AA926959 W76521 W24270 W21526 AA037172 BE267636 H83186 AA469909 N86396 AA001348 BE535736 AA081745 BE556245
 AA082436 H72525 H77575 N49786 W80565 H78746 BE569085 W04339 R98127 T55938 BE279271 AW960304 T29812 AA476873 BE297387
 AA292753 AA177048 NM_001826 X54941 BE314366 AA908783 A719075 BE270172 BE269819 AA863955 A1204530 W25243 A1935150
 AA872039 W72395 T99630 A1422691 H98460 N31428 BE255916 H03265 A1857576 AA776920 AA910644 AA459522 AA293140 AW514567
 R75953 AW662396 AA662522 A1865147 A1423153 AW262230 AA584410 AA583187 AW024595 AW069734 A1828996 AA282997 AA876046
 AW613002 AA527373 AW972459 A1831360 AA621337 AA100926 AA772418 AA594628 A1033892 W95096 A1034317 AA398727 A1085031
 N95210 A1459432 A1041437 AA932124 AA627634 AA935829 A1004827 A1423513 A1094597 H42079 R54703 A1630359 AA617681 AA978045
 AA643280 W44561 A1991988 A1537692 A1090262 AA740817 A1312104 A1911822 AA416871 A1185409 AA129784 AA701623 A1075239
 A1139549 AA633648 A1339996 A1336880 AA399239 A1078708 A1085351 A1362835 A1346618 A1146955 A1989380 A1348243 N92892 AA765850
 A1494230 A1278887 AA962596 A1492600 W80435 AA001979 R97424 A1129015 N24127 AA157451 AA235549 AA459292 AA037114 AA129785
 A1494211 AW059601 AW886710 R92790 N59765 A1361128 AW589407 H47725 H97534 H48076 H48450 T99631 AW300758 H03431 R76789
 AA954344 H77576 R96823 A1457100 N92845 N49682 H42038 BE220698 BE220715 H99552 AA701624 N74173 R54704 H79520 H72923
 H03266 BE261919 AA769633 AA480310 AA507454 AA910586 A1203723 AW104725 W25611 W25071 T88990 H03513 T77589 R99156
 W95095 R97470 AA702275 T77551 AA911952 H82956 N83673 AA283672
 AA009647 AA131254 AA374293 AW954405 H04410 AW606284 AA151166 BE157467 BE157601 H04384 W46291 AW663674 H04021 H01532
 AA190993 H03231 H59605 H01642 AA852876 AA113758 AA626915 AA746952 A1161014 AA099554 R69067

TABLE 14C

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA
 sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 NL_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035, 122804-122921, 124019-124161, 124455-124610, 125672-126076

TABLE 15A: Information for all sequences in Table 16

Table 15A shows the Seq ID No, Pkey, ExAccn, UnigenelD, and Unigene Title for all of the sequences in Table 16.

Table 15B show the accession numbers for those Pkey's lacking UnigenelD's for table 15A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 15C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 15A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Seq ID No: Sequence ID number
Pkey: Unique Eos probeset identifier number
ExAccn: Exemplar Accession number, Genbank accession number
UnigenelD: Unigene number
Unigene Title: Unigene gene title

Seq ID No:	Pkey	ExAccn	UnigenelD	Unigene Title
Seq ID No: 1 & 2	410407	X56839	Hs.63287	carbonic anhydrase IX
Seq ID No: 3 & 4	412719	AW016510	Hs.816	ESTs
Seq ID No: 5 & 6	417034	NM_005183	Hs.80962	neurotensin
Seq ID No: 7 & 8	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam
Seq ID No: 9 & 10	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
Seq ID No: 11 & 12	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
Seq ID No: 13 & 14	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
Seq ID No: 15 & 16	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
Seq ID No: 17 & 18	439285	AL133916		hypothetical protein FLJ20093
Seq ID No: 19 & 20	413753	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kafirin
Seq ID No: 21 & 22	120486	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog
Seq ID No: 23 & 24	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen
Seq ID No: 25 & 26	412140	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabkines
Seq ID No: 27 & 28	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage
Seq ID No: 29 & 30	452838	U65011	Hs.30743	preferentially expressed antigen in mela
Seq ID No: 31 & 32	418663	AK001100	Hs.41690	desmocollin 3
Seq ID No: 33 & 34	418663	AK001100	Hs.41690	desmocollin 3
Seq ID No: 35 & 36	409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibitor
Seq ID No: 37 & 38	429510	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas
Seq ID No: 39 & 40	406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad
Seq ID No: 41 & 42	431846	BE019924	Hs.271580	uroplakin 1B
Seq ID No: 43 & 44	418830	BE513731	Hs.88959	hypothetical protein MGC4816
Seq ID No: 45 & 46	424098	AF077374	Hs.139322	small proline-rich protein 3
Seq ID No: 47 & 48	443648	AI085377	Hs.143610	ESTs
Seq ID No: 49	311034	BE567130	Hs.311389	ESTs, Highly similar to NKGD_HUMAN NKG2-
Seq ID No: 50 & 51	408522	AI541214	Hs.46320	Small proline-rich protein SPRK (human,
Seq ID No: 52 & 53	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
Seq ID No: 54 & 55	435505	AF200492	Hs.211238	interleukin-1 homolog 1
Seq ID No: 56 & 57	417365	BE185289	Hs.1076	small proline-rich protein 1B (comifin)
Seq ID No: 58 & 59	431958	X53629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta
Seq ID No: 60 & 61	441020	W79283	Hs.35962	ESTs
Seq ID No: 62 & 63	423217	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys
Seq ID No: 64 & 65	429538	BE182592	Hs.11261	small proline-rich protein 2A
Seq ID No: 66 & 67	448733	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte
Seq ID No: 68 & 69	444371	BE540274	Hs.239	forkhead box M1
Seq ID No: 70 & 71	444371	BE540274	Hs.239	forkhead box M1
Seq ID No: 72 & 73	444371	BE540274	Hs.239	forkhead box M1
Seq ID No: 74 & 75	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias
Seq ID No: 76 & 77	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias
Seq ID No: 78 & 79	429259	AA420450	Hs.292911	Plakophilin
Seq ID No: 80 & 81	426440	BE382756	Hs.169902	solute carrier family 2 (facilitated glu
Seq ID No: 82 & 83	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an
Seq ID No: 84 & 85	423662	AK001035	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro
Seq ID No: 86 & 87	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino
Seq ID No: 88 & 89	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3
Seq ID No: 90 & 91	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor
Seq ID No: 92 & 93	423634	AW959908	Hs.1690	heparin-binding growth factor binding pr
Seq ID No: 94 & 95	417515	L24203	Hs.82237	ataxia-telangiectasia group D-associated
Seq ID No: 96 & 97	441362	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re
Seq ID No: 98 & 99	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic
Seq ID No: 100 & 101	449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o
Seq ID No: 102 & 103	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3
Seq ID No: 104 & 105	409103	AF251237	Hs.112208	XAGE-1 protein
Seq ID No: 106 & 107	417542	J04129	Hs.82269	progestagen-associated endometrial prote
Seq ID No: 108 & 109	428471	X57348	Hs.184510	stratillin
Seq ID No: 110 & 111	418004	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member
Seq ID No: 112 & 113	414761	AU077228	Hs.77255	enhancer of zeste (Drosophila) homolog 2
Seq ID No: 114 & 115	418203	X54942	Hs.83758	CDC28 protein kinase 2
Seq ID No: 116	447343	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m
Seq ID No: 117 & 118	437016	AU076916	Hs.6398	guanine monophosphate synthetase
Seq ID No: 119 & 120	449230	BE613348	Hs.211579	melanoma cell adhesion molecule
Seq ID No: 121 & 122	446989	AK001898	Hs.16740	hypothetical protein FLJ11036
Seq ID No: 123 & 124	457819	AA057484	Hs.35406	ESTs, Highly similar to unnamed protein
Seq ID No: 125 & 126	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B

	Seq ID No: 127 & 128	414430	AI346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1
	Seq ID No: 129 & 130	418462	BE001596	Hs.85265	integrin, beta 4
	Seq ID No: 131 & 132	100668	L05424	Hs.169610	CD44 antigen (homing function and Indian
5	Seq ID No: 133 & 134	458933	AI638429	Hs.24763	RAN binding protein 1
	Seq ID No: 135 & 136	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 137 & 138	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 139 & 140	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 141 & 142	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
10	Seq ID No: 143 & 144	446269	AW263155	Hs.14559	hypothetical protein FLJ10540
	Seq ID No: 145 & 146	422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur
	Seq ID No: 147 & 148	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con
	Seq ID No: 149 & 150	440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma
	Seq ID No: 151 & 152	439606	W79123	Hs.58561	G protein-coupled receptor 87
15	Seq ID No: 153 & 154	453884	AA355925	Hs.36232	KIAA0185 gene product
	Seq ID No: 155 & 156	453884	AA355925	Hs.36232	KIAA0185 gene product
	Seq ID No: 157 & 158	453884	AA355925	Hs.36232	KIAA0185 gene product
	Seq ID No: 159 & 160	453884	AA355925	Hs.36232	KIAA0185 gene product
	Seq ID No: 161 & 162	404877			NM_005365:Homo sapiens melanoma antigen,
20	Seq ID No: 163 & 164	413129	AF292100	Hs.104613	RP42 homolog
	Seq ID No: 165 & 166	413281	AA861271	Hs.222024	transcription factor BMAL2
	Seq ID No: 167 & 168	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote
	Seq ID No: 169 & 170	416819	U77735	Hs.80205	pim-2 oncogene
	Seq ID No: 171 & 172	451320	AW118072		diacylglycerol kinase, zeta (104kD)
25	Seq ID No: 173 & 174	418543	NM_005329	Hs.85962	hyaluronan synthase 3
	Seq ID No: 175 & 176	454034	NM_000691	Hs.575	aldehyde dehydrogenase 3 family, member
	Seq ID No: 177 & 178	425397	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)
	Seq ID No: 179 & 180	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 181 & 182	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
30	Seq ID No: 183 & 184	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 185 & 186	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 187 & 188	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 189 & 190	419121	AA374372	Hs.89626	parathyroid hormone-like hormone
	Seq ID No: 191 & 192	448993	AI471630	Hs.8127	KIAA0144 gene product
35	Seq ID No: 193 & 194	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR
	Seq ID No: 195 & 196	430393	BE185030	Hs.241305	estrogen-responsive B box protein
	Seq ID No: 197 & 198	425057	AA826434	Hs.1619	achaeta-scutle complex (Drosophila) homol
	Seq ID No: 199 & 200	420462	AF050147	Hs.97932	chondromodulin I precursor
	Seq ID No: 201 & 202	102963	X02404	Hs.274534	calcitonin-related polypeptide, beta
40	Seq ID No: 203 & 204	100576	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid
	Seq ID No: 205 & 206	101175	U82671	Hs.35980	melanoma antigen, family A, 2
	Seq ID No: 207 & 208	429038	AL023513	Hs.194766	seizure related gene 6 (mouse)-like
	Seq ID No: 209 & 210	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
	Seq ID No: 211 & 212	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
45	Seq ID No: 213 & 214	131927	AJ003112	Hs.34780	doublecortin; lissencephaly, X-linked (d
	Seq ID No: 215 & 216	428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN G ANT
	Seq ID No: 217 & 218	427335	AA448542	Hs.251677	G antigen 7B
	Seq ID No: 219 & 220	409420	Z15008	Hs.54451	laminin, gamma 2 (necin (100kD), kalini
	Seq ID No: 221 & 222	114346	AL137256	Hs.130489	ATPase, aminophospholipid transporter-II
50	Seq ID No: 223 & 224	438956	W00847	Hs.135056	Human DNA sequence from clone RP5-850E9
	Seq ID No: 225 & 226	404440			NM_021048:Homo sapiens melanoma antigen,
	Seq ID No: 227 & 228	415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor
	Seq ID No: 229 & 230	103312	Y12642	Hs.3185	lysosomal
55	Seq ID No: 231 & 232	320843	BE069288	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (fr
	Seq ID No: 233	429065	AJ753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT
	Seq ID No: 234 & 235	446102	AW168067	Hs.317694	ESTs
	Seq ID No: 236 & 237	330495	U47924	Hs.71642	guanine nucleotide binding protein (G pr
	Seq ID No: 238	413573	AJ733859	Hs.149089	ESTs
	Seq ID No: 239 & 240	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to
60	Seq ID No: 241 & 242	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to
	Seq ID No: 243 & 244	332180	AF134160	Hs.7327	claudin 1
	Seq ID No: 245	437915	AI637993	Hs.202312	Homo sapiens clone N11 Ntera2D1 teratoca
	Seq ID No: 246 & 247	441553	AA281219	Hs.121296	ESTs
	Seq ID No: 248 & 249	331692	AI683487	Hs.152213	wingless-type MMTV integration site fami
65	Seq ID No: 250 & 251	429413	NM_014058	Hs.201877	DESC1 protein
	Seq ID No: 252 & 253	422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis
	Seq ID No: 254 & 255	448357	N20169	Hs.108923	RAB38, member RAS oncogene family
	Seq ID No: 256 & 257	446292	AF081497	Hs.279682	Rh type C glycoprotein
	Seq ID No: 258 & 259	416209	AA236776	Hs.79078	MAD2 (mitotic arrest deficient, yeast, h
70	Seq ID No: 260 & 261	453922	AF053306	Hs.36708	budding uninhibited by benzimidazoles 1
	Seq ID No: 262 & 263	424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibitor
	Seq ID No: 264 & 265	439223	AW238299	Hs.250618	UL16 binding protein 2
	Seq ID No: 266 & 267	429228	AI553633	Hs.326447	ESTs
	Seq ID No: 268 & 269	409757	NM_001898	Hs.123114	cystatin SN
75	Seq ID No: 270 & 271	411089	AA456454	Hs.214291	cell division cycle 2-like 1 (PITSLRE pr
	Seq ID No: 272 & 273	436511	AA721252	Hs.291502	ESTs
	Seq ID No: 274 & 275	428969	AF120274	Hs.194689	artemin
	Seq ID No: 276 & 277	428969	AF120274	Hs.194689	artemin
	Seq ID No: 278 & 279	428969	AF120274	Hs.194689	artemin
80	Seq ID No: 280 & 281	428969	AF120274	Hs.194689	artemin
	Seq ID No: 282	407137	T97307		gb:ye53h05.s1 Soares fetal liver spleen
	Seq ID No: 283 & 284	412723	AA648459	Hs.335951	hypothetical protein AF301222
	Seq ID No: 285 & 286	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 287 & 288	405770			NM_002362:Homo sapiens melanoma antigen,
85	Seq ID No: 289 & 290	439453	BE264974	Hs.6566	thyroid hormone receptor interactor 13
	Seq ID No: 291 & 292	414774	X02419	Hs.77274	plasminogen activator, urokinase

5	Seq ID No: 293 & 294	424629	M90656	Hs.151393	glutamate-cysteine ligase, catalytic sub
	Seq ID No: 295 & 296	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
	Seq ID No: 297 & 298	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
	Seq ID No: 299 & 300	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
	Seq ID No: 301 & 302	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
	Seq ID No: 303 & 304	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
	Seq ID No: 305 & 306	453968	AA847843	Hs.62711	High mobility group (nonhistone chromoso
	Seq ID No: 307 & 308	403478			NM_022342Homo sapiens kinesin protein 9
10	Seq ID No: 309	441525	AW241867	Hs.127728	ESTs
	Seq ID No: 310 & 311	434105	AW952124	Hs.13094	presentins associated rhomboid-like pro
	Seq ID No: 312 & 313	428810	AF068235	Hs.193788	nitric oxide synthase 2A (inducible, hep
	Seq ID No: 314 & 315	413691	AB023173	Hs.75478	ATPase, Class VI, type 11B
	Seq ID No: 316 & 317	423934	U89995	Hs.159234	forkhead box E1 (thyroid transcription f
15	Seq ID No: 318 & 319	409228	R16811	Hs.22010	ESTs, Weakly similar to 2109260A B cell
	Seq ID No: 320 & 321	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg
	Seq ID No: 322 & 323	413582	AW295647	Hs.71331	hypothetical protein MGC5350
	Seq ID No: 324 & 325	438403	AA806607	Hs.292206	ESTs
	Seq ID No: 326 & 327	403329			unnamed protein product [Homo sapiens]
20	Seq ID No: 328 & 329	408993	AW247090	Hs.57101	minichromosome maintenance deficient (S.
	Seq ID No: 330 & 331	119073	BE245360	Hs.279477	v-ets erythroblastosis virus E26 oncogen
	Seq ID No: 332 & 333	113195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom
	Seq ID No: 334 & 335	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11
	Seq ID No: 336 & 337	101345	NM_005795	Hs.152175	calcitonin receptor-like
25	Seq ID No: 338 & 339	103280	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula
	Seq ID No: 340 & 341	102012	BE259035	Hs.118400	singed (Drosophila)-like (sea urchin fas
	Seq ID No: 342 & 343	105729	H45612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds
	Seq ID No: 344 & 345	134299	AW580939	Hs.97199	complement component C1q receptor
	Seq ID No: 346 & 347	412719	AW016610	Hs.816	ESTs
30	Seq ID No: 348 & 349	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
	Seq ID No: 350 & 351	128924	BE279383	Hs.26557	plakophilin 3
	Seq ID No: 352 & 353	100486	T19006	Hs.10842	RAN, member RAS oncogene family
	Seq ID No: 354 & 355	419121	AA374372	Hs.89626	parathyroid hormone-like hormone
	Seq ID No: 356 & 357	409459	D86407	Hs.54481	low density lipoprotein receptor-related
35	Seq ID No: 358 & 359	330493	M27826		endogenous retroviral protease
	Seq ID No: 360 & 361	417866	AW067903	Hs.82772	collagen, type XI, alpha 1
	Seq ID No: 362 & 363	418113	AJ272141	Hs.83484	SRY (sex determining region Y)-box 4
	Seq ID No: 364 & 365	437016	AU076916	Hs.5398	guanine monophosphate synthetase
	Seq ID No: 366 & 367	429612	AF062649	Hs.252587	pituitary tumor-transforming 1
40	Seq ID No: 368 & 369	440704	M69241	Hs.162	insulin-like growth factor binding prote
	Seq ID No: 370 & 371	431221	AA449015	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye
	Seq ID No: 372 & 373	431565	AF161470	Hs.260622	butyrate-induced transcript 1
	Seq ID No: 374 & 375	431565	AF161470	Hs.260622	butyrate-induced transcript 1
	Seq ID No: 376 & 377	132354	BE185289	Hs.1076	small proline-rich protein 1B (cornlin)
45	Seq ID No: 378 & 379	424441	X14850	Hs.147097	H2A histone family, member X
	Seq ID No: 380 & 381	103768	AF086009	Hs.295398	gbcHomo sapiens full length insert cDNA
	Seq ID No: 382 & 383	417512	X76534	Hs.82226	glycoprotein (transmembrane) nmb
	Seq ID No: 384 & 385	425266	J00077	Hs.155421	alpha-fetoprotein
	Seq ID No: 386 & 387	424503	NM_002205	Hs.149609	integrin, alpha 5 (fibronectin receptor,
50	Seq ID No: 388 & 389	400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
	Seq ID No: 390 & 391	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	Seq ID No: 392 & 393	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	Seq ID No: 394 & 395	418738	AW388633	Hs.6682	solute carrier family 7, (cationic amino
	Seq ID No: 396 & 397	415138	C18356	Hs.295944	tissue factor pathway inhibitor 2
55	Seq ID No: 398 & 399	418506	AA084248	Hs.85339	G protein-coupled receptor 39
	Seq ID No: 400 & 401	423661	D13665	Hs.136348	perlestin (OSF-2os)
	Seq ID No: 402 & 403	414812	X72755	Hs.77367	monokine induced by gamma interferon
	Seq ID No: 404 & 405	417433	BE270266	Hs.82128	ST4 oncofetal trophoblast glycoprotein
	Seq ID No: 406 & 407	417433	BE270266	Hs.82128	ST4 oncofetal trophoblast glycoprotein
60	Seq ID No: 408 & 409	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse
	Seq ID No: 410 & 411	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy
	Seq ID No: 412 & 413	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypothe
	Seq ID No: 414 & 415	400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated
	Seq ID No: 416 & 417	411769	AF245505	Hs.72157	Adican
65	Seq ID No: 418 & 419	428698	AA852773	Hs.334838	KIAA1866 protein
	Seq ID No: 420 & 421	450098	W27249	Hs.8109	hypothetical protein FLJ21080
	Seq ID No: 422 & 423	421552	AF026692	Hs.105700	secreted frizzled-related protein 4
	Seq ID No: 424 & 425	452747	BE153855	Hs.61450	lg superfamily receptor LNIR
	Seq ID No: 426 & 427	450375	AA009647		a disintegrin and metalloproteinase doma
70	Seq ID No: 428 & 429	426215	AW963419	Hs.155223	stanniocalcin 2
	Seq ID No: 430 & 431	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin
	Seq ID No: 432 & 433	432201	AJ538613	Hs.298241	Transmembrane protease, serine 3
	Seq ID No: 434 & 435	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph
	Seq ID No: 436 & 437	442117	AW664964	Hs.128899	ESTs; hypothetical protein for IMAGE:447
	Seq ID No: 438 & 439	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn
75	Seq ID No: 440 & 441	447033	AJ357412	Hs.157601	ESTs
	Seq ID No: 442 & 443	447033	AJ357412	Hs.157601	ESTs
	Seq ID No: 444 & 445	447033	AJ357412	Hs.157601	ESTs
	Seq ID No: 446 & 447	115522	BE614387	Hs.333893	c-Myc target JPO1
	Seq ID No: 448 & 449	410418	D31382	Hs.63325	transmembrane protease, serine 4
80	Seq ID No: 450 & 451	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119
	Seq ID No: 452 & 453	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119
	Seq ID No: 454 & 455	452461	N78223	Hs.108106	transcription factor
	Seq ID No: 456 & 457	412420	AL035668	Hs.73853	bone morphogenetic protein 2
	Seq ID No: 458 & 459	416558	U03272	Hs.79432	fibrillin 2 (congenital contractual ara
85	Seq ID No: 460 & 461	407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antag

	Seq ID No: 462 & 463	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365012.1 [H.s.a
	Seq ID No: 464 & 465	402075			ENSP00000251056*-Plasma membrane calcium
	Seq ID No: 466 & 467	421110	AJ250717	Hs.1355	cathepsin E
5	Seq ID No: 468 & 469	451668	Z43948	Hs.326444	cartilage acidic protein 1
	Seq ID No: 470 & 471	451668	Z43948	Hs.326444	cartilage acidic protein 1
	Seq ID No: 472 & 473	451668	Z43948	Hs.326444	cartilage acidic protein 1
	Seq ID No: 474 & 475	422282	AF019225	Hs.114309	apolipoprotein L
	Seq ID No: 476 & 477	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member
10	Seq ID No: 478 & 479	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),
	Seq ID No: 480 & 481	427747	AW411425	Hs.180655	serine/threonine kinase 12
	Seq ID No: 482 & 483	420281	AI623693	Hs.323494	Predicted cation efflux pump
	Seq ID No: 484 & 485	405932			C15000305:gl 3806122 gb AAC69198.1 (AF0
	Seq ID No: 486 & 487	405932			C15000305:gl 3806122 gb AAC69198.1 (AF0
15	Seq ID No: 488 & 489	444342	NM_014398	Hs.10887	similar to lysosome-associated membrane
	Seq ID No: 490 & 491	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy
	Seq ID No: 492 & 493	417079	U65590	Hs.81134	interleukin 1 receptor antagonist
	Seq ID No: 494 & 495	430890	X54232	Hs.2699	glypican 1
	Seq ID No: 496 & 497	419721	NM_001650	Hs.288650	aquaporin 4
20	Seq ID No: 498 & 499	444471	AB020684	Hs.11217	KIAA0377 protein
	Seq ID No: 500 & 501	413063	AL035737	Hs.75184	chitinase 3-like 1 (cartilage glycoprote
	Seq ID No: 502 & 503	433800	AI034351	Hs.135150	lung type-I cell membrane-associated gly
	Seq ID No: 504 & 505	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro
	Seq ID No: 506 & 507	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro
	Seq ID No: 508 & 509	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitte
25	Seq ID No: 510 & 511	410407	X66839	Hs.63287	carbonic anhydrase IX
	Seq ID No: 512 & 513	309931	AW341683		gb:hd13d01.x1 Soares_NFL_T_GBC_S1 Homo s
	Seq ID No: 514 & 515	412719	AW016610	Hs.816	ESTs
	Seq ID No: 516 & 517	417034	NM_006183	Hs.80962	neurotensin
30	Seq ID No: 518 & 519	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam
	Seq ID No: 520 & 521	413753	U17760	Hs.75517	laminin, beta 3 (nicein (125kD), kalinin
	Seq ID No: 522 & 523	426560	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen
	Seq ID No: 524 & 525	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage
	Seq ID No: 526 & 527	418663	AK001100	Hs.41690	desmocollin 3
35	Seq ID No: 528 & 529	418663	AK001100	Hs.41690	desmocollin 3
	Seq ID No: 530 & 531	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas
	Seq ID No: 532 & 533	406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad
	Seq ID No: 534 & 535	431846	BE019924	Hs.271580	uroplakin 1B
	Seq ID No: 536 & 537	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
40	Seq ID No: 538 & 539	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta
	Seq ID No: 540 & 541	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an
	Seq ID No: 542 & 543	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino
	Seq ID No: 544 & 545	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3
	Seq ID No: 546 & 547	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor
45	Seq ID No: 548 & 549	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3
	Seq ID No: 550 & 551	417542	J04129	Hs.82269	progesterone-associated endometrial prote
	Seq ID No: 552 & 553	449230	BE613348	Hs.211579	melanoma cell adhesion molecule
	Seq ID No: 554 & 555	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma
	Seq ID No: 556 & 557	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma
50	Seq ID No: 558 & 559	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B
	Seq ID No: 560 & 561	418462	BE001596	Hs.85266	integrin, beta 4
	Seq ID No: 562 & 563	410274	AA381807	Hs.61762	hypoxia-inducible protein 2 -
	Seq ID No: 564 & 565	439606	W79123	Hs.58561	G protein-coupled receptor 87
	Seq ID No: 566 & 567	404877			NM_005365: Homo sapiens melanoma antigen,
55	Seq ID No: 568 & 569	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote
	Seq ID No: 570 & 571	418543	NM_005329	Hs.85962	hyaluronan synthase 3
	Seq ID No: 572 & 573	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 574 & 575	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 576 & 577	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
60	Seq ID No: 578 & 579	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 580 & 581	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 582 & 583	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 584 & 585	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR
	Seq ID No: 586 & 587	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
65	Seq ID No: 588 & 589	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
	Seq ID No: 590 & 591	409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalinin
	Seq ID No: 592 & 593	332180	AF134160	Hs.7327	claudin 1
	Seq ID No: 594 & 595	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,
	Seq ID No: 596 & 597	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,
70	Seq ID No: 598 & 599	439223	AWZ38299	Hs.250618	UL16 binding protein 2
	Seq ID No: 600 & 601	409757	NM_001898	Hs.123114	cystatin SN
	Seq ID No: 602 & 603	428969	AF120274	Hs.194689	artemin
	Seq ID No: 604 & 605	428969	AF120274	Hs.194689	artemin
	Seq ID No: 606 & 607	428969	AF120274	Hs.194689	artemin
75	Seq ID No: 608 & 609	428969	AF120274	Hs.194689	artemin
	Seq ID No: 610 & 611	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 612 & 613	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 614 & 615	414774	X02419	Hs.77274	plasminogen activator, urokinase
	Seq ID No: 616 & 617	407944	R34008	Hs.239727	desmocollin 2
80	Seq ID No: 618 & 619	407944	R34008	Hs.239727	desmocollin 2
	Seq ID No: 620 & 621	457489	AU693815	Hs.127179	cryptic gene
	Seq ID No: 622 & 623	429547	AW009166	Hs.99376	ESTs
	Seq ID No: 624 & 625	407242	M18728		gb:Human nonspecific crossreacting antig
	Seq ID No: 626 & 627	407242	M18728		gb:Human nonspecific crossreacting antig
85	Seq ID No: 628 & 629	407242	M18728		gb:Human nonspecific crossreacting antig
	Seq ID No: 630 & 631	444006	BE395085	Hs.10086	type I transmembrane protein Fn14

Seq ID No: 632 & 633	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase domain
Seq ID No: 634 & 635	422109	S73265	Hs.1473	gastrin-releasing peptide
Seq ID No: 636 & 637	419235	AW470411	Hs.288433	neurotrophin
Seq ID No: 638 & 639	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induced
Seq ID No: 640 & 641	419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy
Seq ID No: 642 & 643	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precursor
Seq ID No: 644 & 645	448243	AW369771	Hs.52620	Integrin, beta 8
Seq ID No: 646 & 647	426427	M86699	Hs.169840	TTK protein kinase
Seq ID No: 648 & 649	445537	AJ245671	Hs.12844	EGF-like domain, multiple 6
Seq ID No: 650 & 651	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 6
Seq ID No: 652 & 653	428450	NM_014791	Hs.184339	KIAA0175 gene product
Seq ID No: 654 & 655	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,
Seq ID No: 656 & 657	453392	U23752	Hs.32964	SRY (sex determining region Y)-box 11
Seq ID No: 658 & 659	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic
Seq ID No: 660 & 661	425776	U25128	Hs.159499	parathyroid hormone receptor 2
Seq ID No: 662 & 663	425776	U25128	Hs.159499	parathyroid hormone receptor 2
Seq ID No: 664 & 665	431515	NM_012152	Hs.258583	endothelial differentiation, lysophosphatidylcholine
Seq ID No: 666 & 667	419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7
Seq ID No: 668 & 669	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino acid
Seq ID No: 670 & 671	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino acid
Seq ID No: 672 & 673	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino acid
Seq ID No: 674 & 675	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino acid
Seq ID No: 676 & 677	410001	AB041036	Hs.57771	kallikrein 11
Seq ID No: 678 & 679	426501	AW043782	Hs.293616	ESTs
Seq ID No: 680 & 681	408369	R38438	Hs.182575	solute carrier family 15 (H ⁺ ?? transport)
Seq ID No: 682 & 683	445413	AA151342	Hs.12677	CCG-147 protein
Seq ID No: 684 & 685	422424	AI186431	Hs.296638	prostate differentiation factor
Seq ID No: 686 & 687	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,
Seq ID No: 688 & 689	420610	AI683183	Hs.99348	distal-less homeo box 5

TABLE 15B

Pkey: Unique Eos probe identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT Number	Accession
309931	AW341683	
330493	33264_5	M27826 R78416 AA307645 AW957879 AW957800 AA633529 H03662
439285	47065_1	AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 AJ346341 AI867454 N54784 AI655270 AJ421279 AWO14882
		AA775552 N62351 N59253 AA626243 AI341407 BE175639 AA456968 AI358918 AA457077
450375	83327_1	AA009547 AA131254 AA374293 AW954405 H04410 AW606284 AA151166 BE157467 BE157601 H04384 W46291 AW663674 H04021 H01532
		AA190993 H03231 H59605 H01642 AA852876 AA113758 AA626915 AA746952 AI161014 AA099554 R69067
451320	86576_1	AW118072 AI631982 T15734 AA224195 AI701458 W20198 F26326 AA890570 N90552 AW071907 AI671352 AI375892 T03517 R88265
		AI124088 AA224388 AI084316 AI354686 T33652 AI140719 AI720211 T03490 AI372637 T15415 AW205836 AA630384 T03515 T33230
		AA017131 AA443303 T33623 AI222556 T33511 T33785 AI419606 D55612

TABLE 15C

Pkey: Unique number corresponding to an Eos probe
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham L. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham L. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 NL_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035, 122804-122921, 124019-124161, 124455-124610, 125672-126076
403329	8516120	Plus	96450-96598
403478	9958258	Plus	116458-116564
404440	7528051	Plus	80430-81581
404877	1519284	Plus	1095-2107
405770	2735037	Plus	61057-62075
405932	7767812	Minus	123525-123713

Table 16

Seq ID NO: 1. DNA sequence
Nucleic Acid Accession #: NM_001216
Coding sequence: 43..1422

	1	11	21	31	41	51	
10	GCCCGTACAC	ACCGTGTGCT	GGGACACCCC	ACAGTCAGCC	GCATGGCTCC	CCTGTGCCCC	60
	AGCCCCCTGGC	TCCCTCTGTT	GATCCCGGCC	CCTGCTCCAG	GCCTCACTGT	GCAACTGCTG	120
	CTGTCACTGC	TGCTTCTGAT	GCCTGTCCAT	CCCCAGAGGT	TGCCCCGAT	GCAGGAGGAT	180
	TCCCCCTTGG	GAGGAGGCTC	TTCTGGGAA	GATGACCCAC	TGGCGGAGGA	GGATCTGCCC	240
	AGTGAAGAGG	ATTCACCACG	AGAGGAGGAT	CCACCCGAG	AGGAGGATCT	ACCTGGAGAG	300
15	GAGGATCTAC	CTGGAGAGGA	GGATCTACCT	GAAGTTAAGC	CTAAATCAGA	AGAAGAGGGC	360
	TCCCTGAAGT	TAGAGGATCT	ACCTACTGTT	GAGGCTCCTG	GAGATCCTCA	AGAACCCACG	420
	AATAATGCCC	ACAGGGACAA	AGAAGGGGAT	GACCAGAGTC	ATTGGCGCTA	TGGAGGCGAC	480
	CCGCCCTGGC	CCCGGGTGTG	CCAGCCTGCG	GCGGGCCGCT	TCCAGTCCCC	GGTGGATATC	540
	CGCCCCCAGC	TGCGCCGCTT	CTGCCCCGCC	CTGCGCCCCC	TGGAATCCTC	GGGCTTCCAG	600
20	CTCCCCCGGC	TCCCAAGACT	GCGCCTGCGC	AACAATGGCC	ACAGTGTGCA	ACTGACCCCTG	660
	CCTCTCTGGC	TAGAGATGGC	TCTGGGTCCC	GGGCGGGAGT	ACCGGGCTCT	GCAGCTGCTAT	720
	CTGCATGGG	GGGCTCGAGC	TGCTCGGGC	TGGAGCACA	CTGTGGAAGG	CCACCGTTTC	780
	CCTGCGGAGA	TCCACGTGTT	TCACCTCAGC	ACCGCCTTTG	CCAGAGTTGA	CGAGGCTTGT	840
	GGGCGCCAGC	GAGGCCCTGC	CGTGTGGCC	GCCTTTCTGG	AGGAGGGCCC	GGAGAGAAAC	900
25	AGTGCCTATG	AGCAGTTGCT	GTCTCGCTTG	GAAGAAATCG	CTGAGGAAGG	CTCAGAGACT	960
	CAGGTCCAGT	GACTGGACAT	ATCTGCACTC	CTGCCCTCTG	ACTTCAGCCG	CTACTTCCAA	1020
	TATGAGGGCT	CTCTGACTAC	ACCGCCTGTG	CCCCAGGGTG	TCATCTGGAC	TGTGTTTAACT	1080
	CAGACAGTGA	TGCTGAGTGC	TAAGCAGCTC	CACACCTCTC	CTGACACCTC	GTGGGGACCT	1140
	GGTGACTCTC	GGCTACAGCT	GAACCTCCGA	GCGACGCAGC	CTTTGAATGG	GCGAGTGATT	1200
30	GAGGCTCTCT	TCCCTGCTGG	AGTGGACAGC	AGTCTCTGGG	CTGCTGAGCC	AGTCCAGCTG	1260
	AATTCTCTGC	TGGCTGCTGG	TGACATCCTA	GCCCTGGTTT	TGGCCTCTCT	TTTGTCTGTC	1320
	ACCAGCTGCG	CGTTCTCTGT	GCAGATGAGA	AGGCAGCACA	GAAGGGGAAC	CAAAGGGGGT	1380
	GTGAGCTACC	GCCCCAGAGA	GGTAGCCGAG	ACTGGAGCCT	AGAGGCTGGA	TCTTGAGAGAA	1440
	TGTGAGAAAG	CAGCCAGAGG	CATCTGAGGG	GGAGCCGGTA	ACTGTCTCTG	CTGCTCATT	1500
35	ATGCCACTTC	CTTTAACTG	CCAAGAAATT	TTTTAAATA	AATATTATA	AT	

Seq ID NO: 2 Protein sequence:
Protein Accession #: NP_001207

	1	11	21	31	41	51	
40	MAPLCPSPWL	PLLIPAPAPG	LTVQLLSLL	LLMPVHPQRL	PRMQEDSPLG	GGSSGEDDPL	60
	GEEDLPSEED	SPREDDPGE	EDLPGEEDLP	GEEDLPVQK	KSEEGSLKL	EDLPTVEAPG	120
	DPGEQNNNAH	RDKEGDDQSH	WRYGGDPWP	RVSFACAGRF	QSPVDIRPQL	AAFCPALRPL	180
45	ELLGQLPLPL	PELRRLNNGH	SVQLTLPPL	EMALGPGREY	RALQLHLHW	AAGRPGSEET	240
	VEGHRFPABE	HVVELSTAF	RVDEALRPG	GLAVLAAPLE	EGPEENSAYE	QLLSRLEBIA	300
	EBGSETQVPG	LDISALLPSD	FSRYFYEGS	LTPPCAQGV	IWTVFNTQVM	LSAKQLHTLS	360
	DTLWGPDSR	LQINFRATQP	LNGRVIEASF	PAGVDSPPRA	AEFVQLNSCL	AAGDILALVF	420
	GLLFAVTSVA	FLVQMRQRHR	RGTKGGVSYR	PAEVAETGA			

Seq ID NO: 3 DNA sequence
Nucleic Acid Accession #: BC013923
Coding sequence: 438-1391

	1	11	21	31	41	51	
55	AGCGGGGTTG	TCTATTAAC	TGTTCAAAA	GTATCAGGAG	TTGTCAAGGC	AGAGAAGAGA	60
	GTGTTTGC	AAGGGGGA	GTAGTTTCT	GCCTCTTTAA	GACTAGGACT	GAGAGAAAGA	120
	AGAGGAGAGA	GAAAGAAAG	GAGAGAAAGT	TGAGCCCCAG	GCTTAAGCCT	TTCCAAAAAA	180
60	TAATAATAAC	AATCATCGG	GGCGGCAGGA	TGCGCCAGAG	GAGGAGGAA	GCCTTTTTT	240
	TGATCTGAT	TCCAGTTG	CTCTCTCTT	TTTCCCCCA	AATTATCTT	CGCTGATT	300
	TCCCTGCGGA	GCCTGCGCT	CCGACACCC	CCGCCCGCT	CCCTCTCTC	TCTCCCCCG	360
	CCCGCGGGCC	CCCCAAGTC	CCGCGCGGC	CGAGGGTCG	CGCGCGCGG	CGGGCGGGC	420
	CCGCGCACAG	CGCCGCATG	TACAACATGA	TGGAGACGGA	GCTGAAGCGG	CCGGCGCGC	480
65	AGCAAACTTC	GGGGGCGGC	GGCGGCACT	CCACCGCGC	GGCGGCGGC	GGCAACCAGA	540
	AAAACAGCCC	GGACCGCTC	AAGCGGCCA	TGAATGCCTT	CATGGTGTGG	TCCCGCGGC	600
	AGCGGCGCAA	GATGGCCAG	GAGAACCCCA	AGATGCACAA	CTCGGAGATC	AGCAAGCGCC	660
	TGGGCGCGGA	GTGGAACCTT	TTGTGCGAGA	CGGAGAGCG	GCGTTCATC	GACGAGGCTA	720
	AGCGGCTGCG	AGCGCTGCAC	ATGAAGGAGC	ACCGGATTA	TAAATACCGG	CCCCGGCGGA	780
70	AAACCAAGAC	GCTCATGAAG	AAGGATAAGT	ACACGCTGCC	CGCGGGCTG	CTGGCCCCCG	840
	GCGGCAATAG	CATGGCGAGC	GGGCTCGGG	TGGGCGCGG	CCTGGCGCG	GGCGTGAACC	900
	AGCGCATGGA	CAGTTACCGG	CACATGAACG	GCTGGAGCAA	CGGCAGCTAC	AGCATGATGC	960
	AGGACACAGT	GGGCTACCGG	CAGCACCCCG	GCCTCAATGC	GCACGGCGCA	GCGCAGATGC	1020
	AGCCCATGCA	CGCTACGAC	GTGAGCGCCC	TGCAGTACAA	CTCCATGACC	AGCTCGCAGA	1080
75	CCTACATGAA	CGGCTCGCCC	ACCTACAGCA	TGCTCTACTC	GCAGCAGGGC	ACCCCTGGCA	1140
	TGGCTCTTGG	CTCCATGGGT	TGGTGGTCA	AGTCCGAGGC	CAGCTCCAGC	CCCCCTGTGG	1200
	TTACCTCTTC	CTCCCACTTC	AGGGCGCCCT	GCCAGGCGGG	GGACCTCTGG	GACATGATCA	1260
	GCATGATATC	CCCGGCGGCC	GAGGTGCGGG	AACCGCGCGC	CCCCAGCAGA	CTTCACATGT	1320
	CCAGCATCTA	CAGAGCGCGC	CGGTGCGCG	GCAACGCCAT	TAAACGCACA	CTGCCCCCTC	1380
80	CACACATGTT	AGGGCGGAC	AGCGAACTGG	AGGGGGGAGA	AATTTTCAA	GAAAAACGAG	1440
	GGAAATGGGA	GGGGTGCAAA	AGAGGAGAGT	AAGAAACAGC	ATGGAGAAAA	CCCGGTACGC	1500
	TCAAAAAAAA	AAAAAATAAA	AAAAATCCAT	CACCCACAGC	AAATGACAGC	TGCAAAAGAG	1560
	AAACCAATC	CCATCCACAC	TCACGCAAAA	ACCGCGATGC	CGACAGAAAA	ACTTTTATGA	1620
	GAGAGATCTT	GACTTTCTTT	TKGGGGGACT	ATTTTGTAC	AGAGAAAAAC	TGGGGAGGGT	1680
85	GGGAGGGCGG	GGGGAATGGA	CCTTGTATAG	ATCTGGAGGA	AAGAAAGCTA	CGAAAAACTT	1740
	TTTAAAGGTT	CTAGTGGTAC	GGTAGGAGCT	TTGCAGGAAG	TTTGCAAAAG	TCTTTACCAA	1800
	TAAATATTAG	AGCTAGTCTC	CAAGCGACGA	AAAAATGTTT	TAAATATTGG	CAAGCAACTT	1860
	TGTACAGTA	TTTATCGAGA	TAAACATGGC	AATCAAAATG	TCCATTGTTT	ATAAGCTGAG	1920

AATTGGCCAA TATTTTTCAA GGAGAGGCTT CTTGCTGAAT TTTGATTCTG CAGCTGAAAT 1980
 TTAGGACAGT TGCAACCGTG AAAAGAAGAA AATTATTCAA ATTTGGACAT TTTAATTGTT 2040
 TAAAAATTGT ACAAAGGAA AAAATTAGAA TAAGTACTGG CGAACCATCT CTGTGGTCTT 2100
 GTTTAAAAAG GGCAAAAGTT TTAGACTGTA CTAATTTTAA TAACCTACTG TTTAAAGCAA 2160
 AATGGCCAT GCAGGTTGAC ACCGTTGGTA ATTTATAATA GCTTTTGTTC GATCCCAACT 2220
 TTCCATTTTG TTCAGATAAA AAAAACCATG AAATTACTGT GTTTGAAATA TTTTCTTATG 2280
 GTTGTAAATA TTTCTGTAAA TTTATTGTGA TATTTTAAGG TTTTCCCCCC TTTATTTTCC 2340
 GTAGTTGTAT TTTAAAAGAT TCGGCTCTGT ATTATTGAA TCAGTCTGCC GAGAATCCAT 2400
 GTATATATTT GAACATATAT CATCCTTATA ACAGGTACAT TTTCAACTTA AGTTTTACT 2460
 CCATTATGCA CAGTTTGAGA TAAATAAAT TTTGAAATAT GGACACTGAA AAAAAAAGAA 2520
 AAAAAACAA AACAAAAAA CAAAAACAA AACACAGAAA AACAAAAAA AAAAAACAA 2580
 CACACACAA AACAAAAAA AAAAAAAGA AACAAACAA CAACACAAA CAACACAAA 2640
 CCACACACA AACACACA CACAGAGG

Seq ID NO: 4 Protein sequence:
 Protein Accession #: CAA83435.1

1 11 21 31 41 51
 MYNMMETELK PPGPQQTSGG GGNSTAAAA GGNQKNSPDR VKRPMNAFMV WSRGQRRKMA 60
 QENPKMHNE ISKRLGAEMK LLESETEKRPF IDEAKRLRAL HMKEHPDYKY RPRRKTKTLM 120
 KDKRYTLPGG LLLAPGNSMA SGVGVGAGLG AGVNRMDSY AHMNGWSNGS YSMMDQLGY 180
 PQHPGLNAHG AAQMMPHRY DVSALQYNSM TSSQTYMNGS PTYSMSYSQQ GTPGMALGSM 240
 GSVVKSEASS SPVVTSSSH SRAPCQAGDL RDMISMYLPG AEPPEPAAPS RLHMSQHYQS 300
 GPVPGTAING TPLSEH

Seq ID NO: 5 DNA sequence
 Nucleic Acid Accession #: U91618
 Coding sequence: 29-541

1 11 21 31 41 51
 CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAAATCC AGCTTGTATG 60
 CATGCTACTC CTGGCTTTCA GCTCCTGGAG TCTGTGCTCA GATTGAGAAG AGGAAATGAA 120
 AGCATTAGAA GCAGATTTCT TGACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT 180
 TCCTCTTGGG AAGATGACTC TGCTAAATGT TTGCAGTCTT GTAAATAATT TGAACAGCCC 240
 AGCTGAGGAA ACAGGAGGAG TTCAATGAAGA GGAGCTTGTT GCAAGAAGGA AACTTCTTAC 300
 TGCTTTAGAT GSCTTTAGCT TGGAAAGCAAT GTTGACAATA TACCAGCTCC ACAAATCTG 360
 TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCCTG ATACTGGAAA 420
 TGACAAAAAT GGAAAGGAGG AAGTCATAAA GAGAAAAATT CCTTATATTC TGAACGCGCA 480
 GCTGTATGAG AATAAACCCA GAAGACCCCTA CATACTCAAA AGAGATTCTT ACTATTACTG 540
 AGAGATAAAA TCATTTAATT ACATGTGATT GTGATTCAAT ATCCCTTAAT TAAATATCAA 600
 ATTATATTGT TGTGAAAATG TGACAAACAC ACTTATCTGT CTCTTCTACA ATTGTGGTTT 660
 ATTGAATGTG TTTTCTGCA CTAATAGAAA TTAGACTAAG TGTTTTCAAA TAAATCTAAA 720
 TCTTCAAAA AAAAAAAGAA AATGGGGGCC GCAATT

Seq ID NO: 6 Protein sequence:
 Protein Accession #: AAB50564

1 11 21 31 41 51
 MMAGMKIQLV CMLLLAFSSW SLCSDSEEM KALEADFLTN MHTSKISKAH VPSWKMTLLN 60
 VCSLVNINLS PAEETGEVHE EELVARRKLP TALDGFSLFA MLTYIQLHKI CHSRAPQHEWE 120
 LIQEDILDTG NDKNGKEEVI KRKIPYILKR QLYENKPRRP YILKRDSSYY

Seq ID NO: 7 DNA sequence
 Nucleic Acid Accession #: NM_006536.2
 Coding sequence: 109-2940

1 11 21 31 41 51
 ACCTAAAACC TTGCAAGTTC AGGAAGAAAC CATCTGCATC CATATTGAAA ACCTGACACA 60
 ATGTATGCAG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACAACAT GACCCAAAGG 120
 AGCATTGCAG GTCCCTATTG CAACCTGAAG TTGTGACTC TCCTGGTTCG CTTAAGTTCA 180
 GAACTCCCAT TCCTGGGAGC TGGAGTACAG CTTCAAGACA ATGGGTATAA TGGATTGCTC 240
 ATTGCAATTA ATCCTCAGGT ACCTGAGAAT CAGAACCCTA TCTCAACAT TAAGGAAATG 300
 ATAACCTGAAG CTTCATTTTA CCTATTTAAT GCTACCAAGA GAAGAGTATT TTTGAGAAAT 360
 ATAAAGATTT TAATACCTGC CACATGGAAA GCTAATAATA ACAGCAAAAT AAAACAAGAA 420
 TCATATGAAA AGGCAATGT CATAGTGACT GACTGGTATG GGGCAGATGG AGATGATCCA 480
 TACACCTTAC AATACAGAGG GTGTGGAAAA GAGGGAAAAAT ACATTCAATT CACACCTAAT 540
 TTCCTACTGA ATGATAACTT AACAGCTGGC TACGGATCAC GAGGCGAGT GTTGTGCCAT 600
 GAATGGGCCC ACCTCGGTTG GGGTGTGTTT GATGAGTATA ACAATGACAA ACCTTCTTAC 660
 ATAAATGGGC AAAATCRAAT TAAAGTGACA AGGTGTTTAT CTGACATCAC AGGCATTTTT 720
 GTGTGTGAAA AAGGTCTCTG CCCCCAAGAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA 780
 GGATGCACCT TTATCTACAA TAGCACCCAA AATGCAACTG CATCAATAAT GTTCATGCAA 840
 AGTTTATCTT CTGTGGTTGA ATTTTGTAAAT GCAAGTACCC ACAACCAAGA AGCACCAAAC 900
 CTACAGAACC AGATGTGCAG CCTCAGAAGT GCATGGGATG TAATCAGAGA CTCTGCTGAC 960
 TTTTACCACA GCTTTTCCAT GAATGGGACT GAGCTTCCAC CTCCTCCAC ATTCTGCTT 1020
 GTACAGGCTG GTGACAAAGT GGTCTGTGTT GTGCTGGATG TGTCCAGCAA GATGGCAGAG 1080
 GCTGACAGAG TCCTTCAACT ACAACAAGCC GCAGAAATTT ATTTGATGCA GATTGTTGAA 1140
 ATTCAATCTT TCGTGGCAT TGCCAGTTTC GACAGCAAAG GAGAGATCAG AGCCCACTA 1200
 CACCAAAATTA ACAGCAATGA TGATGGAAG TTGCTGGTTT CATATCTGCC CACCACTGTA 1260
 TCAGCTAAAA CAGACATCAG CATTTGTTCA GGGCTTAAGA AAGGATTGGA GGTGGTTGAA 1320
 AAACGTAATG GAAAGCTTA TGGCTCTGTG ATGATATTAG TGACCAGCGG AGATGATAAG 1380
 CTCTCTGACA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAATTCA CTCCATTGCC 1440

5 CTGGGTTTCAT CTGCAGCCCC AAATCTGGAG GAATTATCAC GTCTTACAGG AGGTTTAAAG 1500
 TTCTTTGTTT CAGATATATC AAACCTCCAAT AGCATGATTG ATGCTTTTCAG TAGAATTTC 1560
 TCTGGAACTG GAGACATTTT CCAGCAACAT ATTCAGCTTG AAAGTACAGG TGAAAATGTC 1620
 AAACCTCACC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT 1680
 ATGTTTCTAG TTACGTGGCA GGCCAGTGGT CCTCTGAGA TTATATTATT TGATCCTGAT 1740
 GGACGAAAAA ACTACACAAA TAATTTTATC ACCAATCTAA CTTTTCGGAC AGCTAGTCTT 1800
 TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCTTGAACAA TACCCATCAT 1860
 TCTCTCAAG CCTGAAAGT GACAGTGACC TCTCGGCCT CCAACTCAGC TGTGCCCCCA 1920
 10 GCCACTGTGG AAGCCTTTGT GGAAGAGAC AGCCTCCATT TTCTCATCC TGTGATGATT 1980
 TATGCCAATG TGAACAGGG ATTATTATCC ATTCTTAATG CCACTGTACG TGCCACAGTT 2040
 GAGCCAGAGA CTGGAGATCC TGTTAOCCTG AGACTCCTTG ATGATGGAGC AGGTGCTGAT 2100
 GTTATAAAAA ATGATGGAAT TTACTOGAGG TATTTTTTCT CCTTTGCTGC AAATGGTAGA 2160
 TATAGCTTGA AAGTGCAATG CAATCACTCT CCCAGCATAA GCACCCCAAC CCACTCTATT 2220
 15 CCAGGGAGTG ATGCTATGTA TGTACCAAGT TACACAGCAA ACGGTAATAT TCAGATGAAT 2280
 GCTCCAAAGA AATCAGAGG CAGAAATGAG GAGGAGCGAA AGTGGGGCTT TAGCCGAGTC 2340
 AGCTCAGGAG GCTCCTTTTC AGTCTGGGA GTTCCAGCTG GCCCCACCC TGATGTGTTT 2400
 CCACCATGCA AAATTATGTA CCTGGAAGCT GTAAAAGTAG AAGAGGAATT GACCCTATCT 2460
 TGGACAGCAG CTGGAGAAGA CTTTGATCAG GGCACGGCTA CAAGCTATGA AATAAGAATG 2520
 20 AGTAAAGTCT TACAGAAATAT CCAAGATGAC TTAAACAATG CTATTTTAGT AAATACATCA 2580
 AAGCGAAATC CTCAGCAAGG TGGCATCAGG GAGATATTGA CGTTCTCACC CCAGATTTCC 2640
 ACGAATGAGC CTGAACATCA GCCAAATGGA GAAACACATG AAGGCCACAG AATTATGTTT 2700
 GCAATACGAG CAATGGATAG GAACTCCTTA CAGTCTGCTG TATCTAACAT TGCCAGGGG 2760
 CCTCTGTTTA TTCCCCCAAA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAAA 2820
 25 GGAATTTTAA CAGCAATGGG TTTGATAGGA ATCAATTGCC TTATTATAGT TGTGACACAT 2880
 CATACTTAA GCAGGAAGAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATTATAA 2940
 ATAAATATCC AAGTGTCTT CCTTCTTAGA TATAAGACCC ATGCCCTTCG ACTACAAAAA 3000
 CATACTAACA AAGTCAAAAT AACATCAAAA CTGTATTAAA ATGCATTGAG TTTTGTGACA 3060
 ATACAGATAA GATTTTATCA TGGTAGATCA ACAATTCITT TTGGGGTAG ATTAGAAAAC 3120
 30 CCTTACACTT TGGCTATGAA CAAATAATAA AAATATTCTT TTAAGTAAT GTCTTTAAAG 3180
 GCAAGGGGAA GCGTAAGAGG GGACCACTGT CAAGGAAGT TTGTTTATT GAGGTGGAAA 3240
 AATAGCCCCA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAACTGTCTG TGTGAAGCAA 3300
 TCATTTAGTT ACTTTGATTA ATTTTCTTCT TCTCCTTATC TGTGCAGTAC AGGTGTCTTG 3360
 TTTACATGAA GATCATGCTA TATTTATAT ATGTAGCCCC TAATGCAAGC CTCTTACCTT 3420
 35 CTGCTATTT TGTATATAT ATTTACAGAT ACATCTCCCT GCTAATGCTC AGAGATCTTT 3480
 TTTCACTGTA AGAGGTAACC TTTAACAATA TGGGTATTAC CTTTGTCTCT TCATACCGGT 3540
 TTTATGACAA AGGTCTATTG AATTATTTTG TGTGTAAGTT TCTACTCCCA TCAAAGCAGC 3600
 TTTCTAGTT TATTGCTTG GGTATTATG GAATGATAGT TATAGCCCN TATAATGCCT 3660
 TACCTAGGAA A

Seq ID NO: 8 Protein sequence:
 Protein Accession #: NP_006527.1

45 1 11 21 31 41 51
 MTQRSIAGPI CNLKFVTLIV ALSSELPFLG AGVQLQDNQY NGLLIAPNP VPENQNLIEN 60
 IKEMITEASF YLEFNATKRRV FFRNIKILIP ATWKANNNSK IKQESYEKAN VIVTDWYGAAH 120
 GDDPYTLQYR GCGKEGKYIH FTFNFIILNDN LTAGYGSRRG VFVHEWAHLR WGVFDEYNND 180
 KPFYINGNQK IKVTRCSSDI TGIFVCEKGP CPQENCIISK LFKEGCTFIY NSTONATASI 240
 50 MFMQSLSSVV EFCNASTINQ EAPNLQNMCM SLRSAMDVIT DSADPHHSFP MNGTELPPPP 300
 TFSLVQAGDK VVCLVLDVSS KMAEADRLQ LQQAEEFYLM QIVRIHTFVG IASFDSKGEI 360
 RAQLHQIINSD DDKRLVSVYL PTTVSAKTDI SICSGLKKGF EVVEKLNGKA YGSVMILVTS 420
 GDDKLLGNCL PTVLSSGSTI HSIALGSSAA PNLEELSRLT GGLKFPVPDI SNSNSMIDAF 480
 SRISSGTGDI FQHQIQLST GENVKPHHQL KNTVTVDNTV GNDTMFLVTW QASGPPEIIL 540
 FDPDRKYYT NNFITNLFFR TASLWIPGTA KPGHWTYTLN NTHSLQALK VTVTSRASNS 600
 55 AVPPATVTEAF VERDLSLHFF PVMIIYANVKQ GYPYILNATV TATVEPETGD PVTLRLLDDG 660
 AGADVIRKNDG IYSRVYFFSA ANGRYSLVKH VNHSPSISTP AHSIPGSHAM YVPGYTANGN 720
 IQMNAPRKSV GRNEEERKVG FSRVSSGGSF SVLGVPAGEH PDVFPCKII DLBAVKVEEE 780
 LTLSTWAPGE DFDQGAQTSY EIRMSKSLQN IQDDFNAIL VNTSKRNPOQ AGIRIEFTFS 840
 60 PQISTNGPEH QPNGETHESH RIYVAIRAMD RNSLQSAVSN IAQAPLFIPP NSDPVPARDY 900
 LILKGVLTAM GLIGIICLII VVTHTLRSRK KRADKKENGK KLL

Seq ID NO: 9 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 336-632

65 1 11 21 31 41 51
 CTCCCCCTCAC CCGGCTCCAG GATGCCCACT CCCCAGGACA CCTCCCACTT CCCACTGTGG 60
 CCTGGGTGGG CTCAGGGGCT GCCCTTGACC TGGCCTAGAG CCTCCCCCA GCTGGTGGTG 120
 GAGCTGGGAC TCTCTGGGAG GGAGGGGGCT GGGAGGGAAT GAGTGGGAAT GGCAAGAGGC 180
 CAGGGTTTGG TGGGATACAA ATCCTCACCC TGGGAGCCTG GCTGCCTTGC TCTCCTTCTT 240
 CCAGTGGGCG CCACATATAA ATCCTCACCC TGGGAGCCTG GCTGCCTTGC TCTCCTTCTT 300
 70 GGGTCTGTCT CTGCCACCTG GTCTGCCACA GATCCATGAT GTGCAGTTCT CTGGAGCAGG 360
 CGCTGGCTGT GCTGCTCACT ACCTTCCACA AGTACTCCTG CCAAGAGGGC GACAAAGTTCA 420
 AGCTGAGTAA GGGGGAATG AAGGAACTTC TGACAAGGA GCTGCCAGC TTTGTGGGGG 480
 AGAAAGTGA TGAGGAGGGG CTGAAGAAGC TGATGGGCAG CCTGGATGAG AACAGTGACC 540
 AGCAGGTGGA CTTCAGGAG TATGCTGTTT TCCTGGCACT CATCACTGTC ATGTGCAATG 600
 ACTTCTTCCA GGGCTGCCCA GACCGACCTT GAAGCAGAAC TCTTGACTTC CTGCCATGGA 660
 80 TCTCTTGGGC CCAGGACTGT TGATGCTCTT GAGTTTGTGA TTCAATAAAC TTTTGTGTC 720
 TGTGTGATAA ATTTTAATTTG CTCACTGATG TTCCATAACC CGGCTGGGCTC AGCTGGAGTG 780
 CTGGGAGATG AGGGCCTCCT GGATCCTGCT CCCTTCTGGG CTCTGACTCT CCTGGAATC 840
 TCTCCAAGGC CAGAGCTATG CTTTAGGTCT CAATTTTGA ATTTCAAACA CCAGCAAAAA 900
 85 ATTGGAATC GAGATAGGTT GCTGACTTTT ATTTGTGCAA ATAAAGATAT TAAAAAGGC 960
 AATAACCA

Seq ID NO: 10 Protein sequence:

Protein Accession #: NP_005969.1

1 11 21 31 41 51
5 MMCSLEQAL AVLVTTFHKY SCQEGDKFKL SKGEMKELLH KELPSFVGEK VDEBGLKKLM 60
GSLDENSQQ VDFQYAVFL ALITVMCNDF FQGCPRDP

Seq ID NO: 11 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 336-626

1 11 21 31 41 51
15 CTCCCTCAC CCGGTCAG GATGCCAGT CCCCACGACA CCTCCCACTT CCCACTGTGG 60
CCTGGGTGGG CTGAGGGCT GCCTTGACC TGGCCTAGAG CCTCCCCCA GCTGGTGGTG 120
GAGCTGGCAC TCTCTGGAG GAGGGGGCT GGGAGGGAAT GAGTGGGAAT GGCAAGAGGC 180
CAGGGTTTGG TGGGATCAG TTGAGGCAGG TTGGTTTCC TTAATAATGCC AAGTGGGGG 240
CCAGTGGGGC CCACATATA ATCTCACC TGGAGCCTG GCTGCCTGC TCTCTTCT 300
GGGTCTGTCT CTGCCACTG GTCTGCCACA GATCCATGAT GTGCAGTTCT CTGAGCAGG 360
CGCTGGCTGT GCTGGTCACT ACCTTCCACA AGTACTCCTG CCAAGAGGGC GACAAGTTCA 420
AGCTGAGTAA GGGGGAATG AAGGAACTTC TGCACAAGGA GCTGCCAGC TTTGTGGGGC 480
ATTCCAGAGA ACCATGTCT GTGAGGGCCT TCCAGTCCA TCTGTTAAT CCTGTCTATTG 540
GAGACTTGAG AAACAGAGC CCAGAAGGA AAAGTGATTG TCCCAAGATC ACACAGCACT 600
GGAGAAAGTG GATGAGGAGG GGCTGAAGAA GCTGATGGGC AGCCTGGATG AGAACAGTGA 660
CCAGCAGCTG GACTTCCAGT AGTATGCTGT TTTCCTGGCA CTGATCAGTG TCATGTGCAA 720
TGACTTCTTC CAGGGCTGCC CAGACCGACC CTGAAGCAGA ACTCTTGACT TCCTGCCATG 780
GATCTCTTGG GCCCAGGACT GTTGATGCCT TTGAGTTTGG TATTCAATAA ACTTTTITG 840
TCTGTTGATA ATATTTAAT TGCTCAGTGA TGTTCCATAA CCGGCTGGC TCAGCTGGAG 900
TGCTGGGAGA TGAGGGCCTC CTGGATCCTG CTCCCTCTG GGCCTGACT CTCTGGA 960
TCTCTCCAG GCCAGAGCTA TGCTTTAGGT CTCATTTTGG GAATTTCAA CACCAGCAA 1020
AAATTGAAA TCGAGATAGG TTGCTGACTT TTATTTTGTG AAATAAAGAT ATTAATAAAG 1080
GCAATACCA

Seq ID NO: 12 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
40 MMCSLEQAL AVLVTTFHKY SCQEGDKFKL SKGEMKELLH KELPSFVGHs REPCA VRAFR 60
VHLFPNPVIGD LRNQSPEGKS DCPKITQHWR KWMRRG

Seq ID NO: 13 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 58-354

1 11 21 31 41 51
50 GTGAGCTCAC CATGTGGGGG TGAGGCTGAG AGAAAACAAG TACACAGCCA CAGATCCATG 60
ATGTGCACTT CTCTGGAGCA GGGCTGGCT GTGCTGGTCA CTACCTTCCA CAAGTACTCC 120
TGCAAGAGG GGCACAAGTT CAAGCTGAGT AAGGGGGAAT TGAAGGAAT TCTGCACAAG 180
GAGCTGCCCA GCTTTGTGGG GGAGAAAGTG GATGAGGAGG GGCCTGAAGAA GCTGATGGGC 240
AGCCTGGATG AGAACAGTGA CCAGCAGGTG GACTTCCAGG AGTATGCTGT TTTCTGGCA 300
55 CTCATCAGTG TCATGTGCAA TGACTTCTTC CAGGCTGCC CAGACCGACC CTGAAGCAGA 360
ACTCTTGACT TCCTGCCATG GATCTCTTGG GCCCAGGACT GTTGATGCCT TTGAGTTTGG 420
TATTCAATAA ACTTTTITG TCTGTTGATA ATATTTAAT TGCTCAGTGA TGTTCCATAA 480
CCCGCTGGC TCAGCTGGAG TGCTGGGAGA TGAGGGCCTC CTGGATCCTG CTCCCTCTG 540
GGCTCTGACT CTCCTGAAA TCTCTCCAAG GCCAGAGCTA TGCTTTAGGT CTCATTTTGG 600
60 GAATTTCAA CACCAGCAA AAATTGAAA TCGAGATAGG TTGCTGACTT TTATTTTGTG 660
AAATAAAGAT ATTAATAAAG GCAATACCA

Seq ID NO: 14 Protein sequence:
Protein Accession #: NP_005969.1

1 11 21 31 41 51
65 MMCSLEQAL AVLVTTFHKY SCQEGDKFKL SKGEMKELLH KELPSFVGEK VDEBGLKKLM 60
GSLDENSQQ VDFQYAVFL ALITVMCNDF FQGCPRDP

Seq ID NO: 15 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 62-358

1 11 21 31 41 51
75 GGAGGGTGTG CCGCTGAGT ACTGCTGGG CATCTGGGCC TGGAACTCG GCCACAGATC 60
CATGATGTGC AGTTCTCTGG AGCAGGCGCT GGCTGTGCTG GTCACTACCT TCCACAAGTA 120
CTCCTGCCAA GAGGGCAGCA AGTTCAAGCT GAGTAAGGGG GAAATGAAGG AACTCTGCA 180
80 CAAGGAGCTG CCCAGCTTGG TGGGGAGAA AGTGGATGAG GAGGGGCTGA AGAAGCTGAT 240
GGGAGCCTG GATGAGAACCA GTGACCAGCA GGTGGACTTC CAGGAGTATG CTGTTTCTCT 300
GGCATCATC ACTGTCTATG GCAATGACTT CTTCAGGGC TGCCAGAGCC GACCTGAAG 360
CAGAATCTT GACTTCTGCG CATGGATCTC TTGGGCCAG GACTGTGAT GCCTTGTAGT 420
75 TTTGTATTCA ATAACTTTT TTTGCTGTT GATAATATT TAATTGCTCA GTGATGTTCC 480
ATAACCCGGC TGGCTCAGCT GGAGTGTCTG GAGATGAGGG CCTCCTGGAT CTGCTCCCT 540
TCTGGGCTCT GACTCTCTG GAAATCTCTC CAGGCCAGAG GCTATGCTTT AGGTCTCAAT 600
TTTGAATTT CAAACACCA CAAAAAATTG GAAATGAGAG TAGGTGCTG ACTTTTATTT 660

Seq ID NO: 16 Protein sequence:
Protein Accession #: NP_005969.1

5
1 11 21 31 41 51
MMCSLEQAL AVLVTFTHKY SCQEGDKFKL SKGEMKELLE KELPSFVGEK VDEBGLKKLM 60
10 GSLDENSDDQ VDFQEYAVFL ALITVMCNDP FQGCDFRP

Seq ID NO: 17 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 939-2372

15
1 11 21 31 41 51
AAGACGGATT CTCAGACAAG GCTTGCAAAAT GCCCCGCGAG CATCATTTAA CTGCACCGCG 60
20 AGAATAGTTA CGGTTTGTC CCGACCCCTC CCGGATCGCC TAATTTGTCC CTAGTGAGAC 120
CCGAGGCTCT TGCCCGCGCC TGGCTTCTTC GTAGCTGGAT GCATATCGTG CTCGGGCGAG 180
CGCGGGCGCA GGGCAGCGGT TCGGCGACAC CCTAGCACAC ATGAACACGC GCAAGAGCTG 240
AACCAAGCAC GGTTTCCATT TCAAAAAGGG AGACAGCCTC TACCGCGATT GTAGAAGAGA 300
CTGTGGTGTG AATTAGGAGC CGGGAGGCGT CGAACGAGG AACGGTTCAT CTTAGAGACT 360
25 AATTTTCTGG AGTTTCTGCC CCTGCTCTGC GTACGCCCCC ACGTCACTTC GCCAGCAGTA 420
GCAGAGCGCG CGGCGGCGCG TCCCGGAATT GGGTTGGAGC AGGAGCCTCG CTGGCTGCTT 480
CGCTCGCGCT CTAAGCGCTC AGTCCCGCGC GGTAGCAGGA GCCTGGACCC AGGCGCGCGC 540
GGCGGGCGTG AGGCGCGCGA GCCCGGCGCT CAGGTGCATA CCGGACCCCG ATTGCGCATCT 600
30 AACCAAGGAT CTGCGCCCCA GAGAGTCCCG GGAGCGCGCG CGGTCTGGTG CCGGCGCGCC 660
GGGCCATGCA CGACAGCGCG CCGCGGAGCT GTAGCGCCCC CCTGTAAAGC 720
GGTTGCGTAT GCCCGGGGCA CTGTGAACCC TGCCGCTGCG CGGAACACTC TCGCTCCGG 780
ACCAAGTCAG CCTCTGATAA GCTGGACTCG GCACGCGCGC AACCAAGCAC GAGGAGTTAA 840
GAGAGCGGCA AGGCGAGGGA AGGCTCCCGC GCACGGGTGG GGAAGCGCG CCGGTGCAGC 900
35 GCGGGGAGAC GCACTCGGCG TGGCACTGGC TGCTAGGATG GTCGTCTCTG ATAAGGTGGC 960
ATGACCGCGC CATGGCGCGG CTCTGGGGCT TCTGCTGGCT GGTGTGGGCG TTCTGGAGGG 1020
CCGCTTTCGC CTGTCCACAG TCCTGCAAAAT GCAGTGCCTC TCGGATCTGG TGCAGCGACC 1080
CTTCTGCTAG ATGCTGCGCA TTTCGAGAT TGGAGCCTAA CAGTGTAGAT CCTGAGAACTA 1140
TCACCGAAAT TTTCATGCGA AACCAAGAAA GGTAGAAAT CATCAAGCA GATGATGTTG 1200
40 AAGCTTATGT GGGAGTGAAG AATCTGACAA TTGTGGATTC TGGATTAAAA TTGTGGCTC 1260
ATAAAGCATT TCTGAAAAAC AGCAACCTGC AGCACATCAA TTTTACCGA AACCAACTGA 1320
CGAGTTTGTG TAGGAACATG TTCCGTCACC TTGACTTGTC TGAATGATC CTGGTGGGCA 1380
ATCCATTATG ATGCTGCTGT GACATTATGT GGATCAAGAC TCTCAAGAG GCTAAATCCA 1440
GTCCAGACAC TCAGGATTGG TACTGCCTGA ATGAAGCAG CAAGAATATT CCGCTGGCAA 1500
45 ACCTGCAGAT ACCCAATTGT GGTTCGCCAT CTGCAAAATC GGCGCGCACT AACCTCACTG 1560
TGGAGGAAGG AAGGTCTATC ACATTATCCT GTAGTGGGCG AGGTGATCGG GTTCCTAATA 1620
TGTAATTGGA TGTGTGTAAC CTGGTTTCCA AACATATGAA TGAACAAGC CACACACAGG 1680
GCTCCTTAAG GATAACTAAC ATTTCTCCG ATGACAGTGG GAAGCAGATC TCTTGTGTGG 1740
CGGAAATCTT TGTAGAGGAA GATCAAGATT CTGTCAACCT CACTGTGATC TTGCAACCAA 1800
50 CTATCACTT TCTCGAATCT CCAACCTCAG ACCACCACTG GTGCATTCCA TCACTGTGA 1860
AAGGCAACCC CAACACAGCG CTTCAGTGGT TCTATAACGG GGCAATATTG AATGAGTCCA 1920
AATACATCTG TACTAAATTA CATGTTACCA ATCACACGGA GTACACCGCG TGCTCCAGC 1980
TGGATAATCC CACTCACTAG AACCAATGGG ACTACACTCT AATAGCCAG AATGAGTATG 2040
GGAAGGATGA GAAACAGATT TCTGCTCACT TCATGGGCTG GCCTGGAAAT GACGATGGTG 2100
55 CAACCCCAA TTATCTGAT GTAATTATG AAGATTATGG AACTGCAGCG AATGACATCG 2160
GGGACACCA GACAGAGAGT AATGAATACC CTTCACAGA CGTCACTGAT AAAACCGGTC 2220
GGGAACATCT CTGGCTCTAT GCTGTGGTGG TGATTGCGTC TGTGGTGGGA TTTTGCTTTT 2280
TGGTAATGCT GTTTCGCTAT AAGTTGGCAA GACACTCCAA GTTTGGCATG AAAGGTTTTT 2340
60 TTTTGTTCCT TAAGATCCCA CTGGATGGGT AGCTGAAATA AAGGAAAGA CAGAGAAAGG 2400
GGCTGTGGTG CTGTGTGGTT GATGCTGCCA TGTAAAGCTG ACTCCTGGGA CTGCTGTGG 2460
CTTATCCCGG GAAGTGTCTT TTATCTGGGG TTTTCTGGTA GATGTGGGCG GTGTTTGGAG 2520
GCTGTACTAT ATGAAGCCTG CATATACTGT GAGCTGTGAT TGGGGAACAC CAATGCAGAG 2580
GTAACCTCA GGCAGCTAAG CAGCACTCA AGAAAAATG TTAATAATAT GCTTCTCTTC 2640
TTACAGTAGT TCAATACAA AACTGAAATG AAATCCCATG GATTTGACT TCTCTTCTGA 2700
65 AAAGTGTGCT TTTTGACCCT ACTGGACAT TATTGACTTA ATTGCTTCTG TTTATTAAAA 2760
TTGACCTGCA AAGTTAAAAA AAAATTAAAG TTGAGAACAG GTATAAGTGC ACACCTGAATA 2820
GTCTAATCTA CATGTAACAC ATATTTTATG GTGATTTTCT ATACTCTAAT CAGCACTGAA 2880
TTCAGAGGGT TTGACTTTTT CATCTATAAC ACAGTGACTA AAAGAGTTAA GGGTATATAT 2940
ACCATCACTT TGGGACTTGG TAGTATTATT AAAAGGTTAT TTCCTTCACT GTCAATAAAA 3000
70 GTCCAAATGT TTAGCTTAGG TCTGAGAGTC AAACAATGTT AAGGATTGTC TTAAGTTTCC 3060
TTAGCCAGCA AAACAAAACA AAACAAAACA AACCAATGAA AACCGTTTAA AAAGAAGAAG 3120
AAGAAAAAAA ACAGAACAA GCAGCAACAG CTGTTTTGTT GGGGCTATAG ATTTAAGTTA 3180
GGCATACTCA ATTTAGAAAT AACTAAGAGT GGAATATATG CATATGGTGA AATTATAACC 3240
TTGCCCTTTT TTATTTGCCC TCTGCGATCC ACCTGCTTTT TAGAAGTCTG CCGAGTGAGA 3300
75 AGGCCACAGT ATCTCATGCT GTTTGCATTA CAGAACTGCA GCTTTTCTAC TCTGAAAAGG 3360
CCTGGGAGCA GAATGGCTGG CCTGCTGTGA GCAGGAGAGG AGATTCTAAG AAGGATAGTC 3420
CCCCCTACAA CATACTGTCA TACTGCTGGG TTTTCTATGG TAGGAAAGCT TGTCTTGACC 3480
CCAGCAGCAA AGAGGTGGCA GGTGCTAAT GAATATATGC TTTATAATGT CTTCTTCTAT 3540
TGCTGAGAGG GCAGCCTTAG AGCTGTGGAT TTCTGCATCC CCGCTGAGTC TGACCCATGG 3600
ACACCTGTTT CATTCACCTT AGCATCACAG TGACCTTTGT ATGCTCTGTT CAGTCTGTGT 3660
80 CAGGCACTAT GCTTGTCTTG AAGAGAGGTT TGGCTATCCC CACCCACCCC CACCCACCCC 3720
TGTTCTCTTT TTATCAGGAG GACTTCAGAG CCAGGCTGCG AGCATTTTGT TTGAAAACAC 3780
AATCAGCTCT GACAGTACA CATGCACACA GACGCCATAG CTGAGTTGGA AACATTGATG 3840
TTTTAAAAAT TTATTTTTTT TGGAAATAGT TGCACAAATG CTGCAATTTA GCTTTAAGGT 3900
TCTATAGATT TTTAAGTAGT CCAACACAGT CAGAAACATT GTTTTGAATC CTCTGTAAC 3960
85 CAAGGCATTA ATCTTAATAA ACCAGGATCC ATTTAGGTAC CACTTGATAT AAAAAGGATA 4020
TCCATAATGA ATATTTTATA CTGCATCCTT TACATTAGCC ACTAAATACG TTATTGCTTG 4080
ATGAAGACCT TTCACAGAAT CCTATGGATT GCAGCATTTT ACTTGGCTAC TTCATACCCA 4140

5
10

```

TGCTTAAAG AGGGGAGATT TCTCAAAAGC AGAAACATGC CGCCAGTTCT CAAGTTTTC 4200
TCCTAACTCC ATTGAATGT AAGGGCAGCT GGCCCCCAAT GTGGGGAGGT CCGAACATT 4260
TCTGAATTCC CATTTCTCTG TTGCGGCTA AATGACAGTT TCTGTCAATTA CTTAGATTCC 4320
GATCTTTCCC AAAGGTGTG ATTACAAAG AGGCCAGCTA ATAGCAGAAA TCATGACCC 4380
GAAAGAGAGA TGAATTTCAA GCTGTGAGCC AGGCAGGAGC TCAGTATGGC AAAGGTTCTT 4440
GAGAATCAGC CATTTGGTAC AAAAAAGATT TTTAAAGCTT TTATGTTATA CCATGGAGCC 4500
ATAGAAAGGC TATGGATTGT TTAAGAACTA TTTAAAGTG TTCCAGACCC AAAAAGGAAA 4560
AATAAAAAAA AAGGAATATT TGTACCCAAC AGCTAGAAGG ATTGCAAGGT AGATTTTGT 4620
TTTAAATGCG AGAGAAGTGG ACAGATAAGG CCATTTAATA TATCAAAGAT CAGTTGACAT 4680
CTCCTAGGGA ATGATGAAA CAGCAGGCTA T

```

Seq ID NO: 18 Protein sequence:
Protein Accession #: CAAS3571

15
20
25

```

1 11 21 31 41 51
MSSWIRNHWP AMARLWGFCE LVVGFWRAAF ACPTSCCKSA SRWCSDPSP GIVAFPRLEP 60
NSVDPEINTE IFIANQKRL E INEDDVEAY VGLRNLTIIV SGLKFAHKA FLKNSHLQHI 120
NFRNKLTL SRKHFRLDL SELILVGNPF TCSCDIMWIK TLQEAQSPD TDLYCLNES 180
SKNIFLANLQ IPNCGLP SAN LAAPNLTV E GKSTLSCSV AGDPVPMYV DVGNLVSKHM 240
NETSHQSL RITNISSDD SKQISCVAEN LVGEDQDSV LTVHFAPTIT FLESPTSDH 300
WCIPFTVKGN PKPALQWFPY GAILNESKYI CTKIHVNTHT EYHGCLQLDN PTHMNGDYT 360
LIAKNEYKGD EKQISAHFMG WPGIDGANG NYPDVIYEDY GTAANDIGDT TNRSNEIPST 420
DVTDKTGREH LSVYAVVVIA SVVGFCLLV LFLKLRLHS KFGMKGFVLP HKIPLDG

```

Seq ID NO: 19 DNA sequence
Nucleic Acid Accession #: NM_000228
Coding sequence: 82-3600

30
35
40
45
50
55
60
65
70
75
80
85

```

1 11 21 31 41 51
GCCTTCAGGC GATCTGGAGA AAGAAGCGCA GAACACACAG CAAGGAAAGG TCCTTCTGG 60
GGATCAACCC ATTGGCTGAA GATGAGACCA TTCTTCTCT TGTGTTTTC CCTGCTGGC 120
CTCTGCTCATG CCCAACAAAGC CTGCTCCCGT GGGGCTCTGT ATCCACCTGT TGGGGACCTG 180
CTTGTGTGGGA GGAGCCGGTT TCTCCGAGCT TCATCTACCT GTGGACTGAC CAAGCTGAG 240
ACCTACTGCA CCAGTATGG CGAGTGGCAG ATGAAATGCT GCAAGTGTGA CTCAGGCGC 300
CCTCAAACT ACTACAGTCA CGAGTAGAG AATGTGGCTT CATCCTCCGG CCCCATGGC 360
TGTGGCAGT CCAGAAATGA TGTGAACCTT GTCTCTCTG AGCTGGACCT GGACAGGAGA 420
TTCCAGCTTC AAGAAGTCAT GATGGAGTTC CAGGGGCCCA TGCCCGCGG CATGCTGATT 480
GAGCGCTCCT CAGACTTCGG TAAGACCTGG CGAGTGTACC AGTACCTGGC TGCCGACTGC 540
ACCTCCACCT TCCTCCGGGT CGGCCAGGCT CGGCTCTAGA GCTGGCAGGA TGTTCGGTTC 600
CAGTCCCTGC CTCAGAGGCG TAATGCACGC CTAATGGGG GGAAGGTCCA ACTTAACCTT 660
ATGATTTAG TGTCTGGGAT TCCAGCAACT CAAAGTCAA AAATTCAGA GGTGGGGGAG 720
ATCACAACT TGAGAGTCAA TTTACCCAGG CTGGCCCTCT TGCCCCAAG GGGCTACCAC 780
CTTCCAGCG CCTACTATGC TGTGTCCAG CTCGTCTGC AGGGGAGCT CTCTGTGAC 840
GGCCATCTG ATCGCTGGC ACCCAAGCT GGGGCTCTG CAGGCCCTC CACCGCTGTG 900
CAGGTCCAG ATGTCTGTGT CTGCCAGCAC AACACTGCCG GCCCAAATG TGAGCGCTGT 960
GCACCTCTT ACACACACCG GCCCTGGAGA CCGCGGAGG GCCAGGACG CCAATGAATG 1020
CAAAGGTGG ACTGCAATGG GCACTCAGAG ACATGTCACT TTGACCCCG TGTGTTGCC 1080
GCCAGCCAG GGCATATGG AGGTGTGTGT GACAATTGCC GGCACACAC CGAAGGCAAG 1140
AACTGTGAGC GGTGTGAGT GCACTATTTC CGGAACCGG GCCCGGAGC TTCCATTGAC 1200
GAGACCTGAG TGTCTGCGA GTGTGATCCG GATGGGGCAG TGCCAGGGG CTCCCTGTGAG 1260
CCAGTGACCG GGCAGTGTGT GTGCAAGGAG CATGTGCAGG GAGAGCGCTG TGACCTATGC 1320
AAGCGGGCT TCACTGGACT CACCTACGCC AACCCGAGG GCTGCCACG CTGTGACTGC 1380
AACATCTCG GGTCCCGGAG GGACATGCCG TGTGACGAG AGATGGGGC CTGCTTTGT 1440
CTGCCCAAG TGGTGGGTCC CAAATGTGAC CAGTGTGCTC CTATCCACTG GAAGCTGGCC 1500
AGTGGCCAG GTGTGAAAC GTGTGCTGC GACCCGACA ACTCCCTCA GCCCACTG 1560
CAACCAATTC ACAGGCGAGT GCCCTGTGG GAAGGCTTT GTGGCTGAT GTGCAAGCT 1620
GCAGCCATTC GCCAGTGTCC AGACCGGACC TATGGAGACG TGGCCACAG ATGCCAGGCC 1680
TGTGACTGTC ATTTCCGGGG AACAGAGGGC CCGGGCTGG ACAAGGCATC AGGCGCTGC 1740
CTCTGCCGCC CTGGCTTGAC CGGCCCCCG TGTGACCACT GCCACGAGG CTACTGCAAT 1800
CGCTACCGC TGTGCTGGC CTGCCACCTT TGCTTCCAGA CCTATGATG GGACCTCCGG 1860
GAGCAGGCC TGCGCTTGG TAGACTCCGC AATGCCACCG CCAGCTGTG GTGAGGCT 1920
GGGCTGGAG ACCGTGGCCT GGCCTCCCG ATCTAGATG CAAAGAGTAA GATTGAGCAG 1980
ATCCGAGCAG TTCTCAGCAG CCCCAGCTC ACAGAGCAGG AGGTGGCTCA GGTGGCCAGT 2040
GCCATCTCT CCTCAGCGC AACTCTCCAG GGCCTGCAGC TGGATCTGCC CTGGAGGAG 2100
GAGAAGTTGT PCCCTCCGAG AGACCTGGAG AGTCTTGACA GAAGCTTCAA TGGTCTCCTT 2160
ACTATGTATC AGAGGAAGAG GGAGCAGTTT GAAAAATAA GCAGTGTGA TCCTTCAGGA 2220
GCCTTCCGA TGCTGAGCAG AGCCTACGAG CAGTCAGCCC AGGCTGCTCA GCAGGTCTCC 2280
GACAGCTGCG GCCTTTTGA CCAGCTCAGG GACAGCCGGA GAGAGGCAGA GAGGCTGGT 2340
CGGCAAGCG GAGGAGGAG AGGCACCGG AGCCCCAAGC TTGTGGCCCT GAGGCTGGAG 2400
ATGTCTTGT TGCTGACCT GACACCCACC TTCAACAAGC TCTGTGGCAA CTCAGGCGAG 2460
ATGGCTGCA CCCCATATC ATGCCCTGGT GAGCTATGTC CCAAGACAA TGGCACAGC 2520
TGTGGCTCCC GCTGAGGGG TGTCTTCCC AGGCGCGGT GGGCTTCTT GATGGCGGG 2580
CAGGTGGCTG ACAGCTGGC GGGCTTCAAT GCCAGCTCC AGCGGACCA GCAGATGATT 2640
AGGGCAGCG AGGAATCTGC CTCACAGATT CAATCCAGT CCCAGCGCTT GGAGACCCAG 2700
GTGAGCGCCA CGCGCTCCA GATGGAGGAA GATGTGAGC GCACACGGCT CCAATCCAG 2760
CAGGTCCGG ACTTCTTAAC AGACCCGAC ACTGATGCAG CCACTATCCA GGAGGTGAGC 2820
GAGGCGTGC TGGCCCTGT GCTGCCACA GACTCAGCTA CTGTTCTGCA GAAGATGAAT 2880
GAGATCCAG CATTTGAGC CAGGCTCCCC AACGTGAGT TGGTCTGTC CCAGACCAAG 2940
CAGGACATTG CGCGTGGCG CGGTTGCAG GCTGAGGCTG AGGAAGCCAG GAGCGAGCC 3000
CATGCAGTG AGGGCAGGT GGAAGATGTG GTTGGGAACC TGCGGAGGG GACAGTGGCA 3060
CTGCAGGAG CTCAGGACAC CATGCAAGG ACCAGCGCT CCTTCCGCT TATCCAGGAC 3120
AGGTTGCTG AGTTTCAGCA GGTACTCGG CCAGCAGAAA AGCTGGTGAC AAGCATGACC 3180
AAGCAGCTG GTGACTCTG GACACGGATG GAGGAGCTCC GCCACCAAG CCGGACGAGC 3240
GGGGCAGAG CAGTCCAGG CCAGCAGCTT GCGGAAGGTG CCAGCGAGCA GGCATTGAGT 3300
CCCCAGAGG GATTGAGAG AATAAACAA AAGTATGCT AGTTGAAGGA CCGTTGGGT 3360

```

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

CAGAGTTCCA TGCTGGGTGA GCAGGGTGCC CGGATCCAGA GTGTGAAGAC AGAGGCAGAG 3420
GAGCTGTTTG GGGAGACCAT GGAGATGATG GACAGGATGA AAGACATGGA GTTGGAGCTG 3480
CTGCGGGCCA GCCAGGCCAT CATGCTGGCC TCGGCGGACC TGACAGGACT GGAGAAGCGT 3540
GTGGAGCAGA TCCGTGACCA CATCAATGGG CGCGTGCTCT ACTATGCCAC CTGCAAGTGA 3600
TGCTACAGCT TCCAGCCCGT TGCCCCACTC ATCTGCGGCC TTTGCTTTTG GTTGGGGGCA 3660
GATTGGGTGT GAATGCTTTC CATCTCCAGS AGACTTTTCA GCAGCCTAAA GTACAGCCTG 3720
GACCACCCCT GGTGTGTAGC TAGTAAGATT ACCCTGAGCT GCAGCTGAGC CTGAGCCAAT 3780
GGGACAGTTA CACTTGACAG ACAAGATGG TGGAGATTGG CATGCCATTG AAACATAAGAG 3840
CTCTCAAGTC AAGGAAGCTG GGCTGGGCAG TATCCCCCGC CTTTAGTTCT CCATCTGGGA 3900
GGAATCCTGG ACCAAGCACA AAAACTTAAC AAAAGTGATG TAAAAATGAA AAGCCAAATA 3960
AAAACTTTG G

Seq ID NO: 20 Protein sequence:
Protein Accession #: NP_000219

1 11 21 31 41 51
MRPFFLLCPA LPGLLHAQQA CSRGACYPPV GDLLVGRTRF LRASSTOGLT KPETYCTQYG 60
EWQMKCKCD SRQPHNYSH RVENVASSSG PMRWQSQND VNFVSLQLDL DRRFQLQEVN 120
MEFGGMPAG MLIERSSDFG KTWVYQYLA ADCTSTFPRV RQGRPQSQND VRCQSLPQRP 180
NARLNGGKVQ LNLMDLVSGI PATQSQKIQE VGEITNLRVN PTRLAPVPQR GYHPFSAYYA 240
VSQRLRLQSSC FCHGHADRCR PKPGASAGPS TAVQVHDCV CQHNTAGPNC ERCAFFYNRR 300
PWRPAEQDA HECQRQDCNG HSSTCHFDPA VFAASQGAYG GVCNDCRDHT EGKNCERCQL 360
HYFNRRRPA SIQETCISCE CDPDGAVPGA PCDPVTGQCV CKEHVQGERC DLCKPGFTGL 420
TYANPQGRHR CDNLLGSRD DMPDEESGR CLCLPNVVGK KCDQCAPYHW KLASGQCEP 480
CACDPHNSPQ PTVPQVHRAV PCREGFGGLM CSAAIIRQCE DRTYGDVATG CRACDCDFRG 540
TEGPGCDKAS GRCLCRPGLT GPRCDQQRG YCNRYFVCA CHPCFQTYDA DLREQALRF 600
RLRNATASLW SGPGLEDRLG ASRILDASK IEQIRAVLSS PAVTEQEVAG VASAILSLRR 660
TLQGLQLDLF LEBETLSLPR DLESILDRSFN GLLTMYQRKR EQFEKISSAD PSGAFRMLST 720
AYEQSAQAQ CDNLSRLLD QLRDSREAE RLVRQAGGGG GTGSPKLVAL RLEMSSLPDL 780
TPTFNKLCGN SRQMACTPIS CPGELCPQDN GTACGSRCRG VLPAGGAFI MAGQVAEQLR 840
GFNAQLQRTQ OMIRAAEESA SIQSSAQRL ETQVSASRSQ MEEDVRRTRL LIQVDRFLT 900
DPDTRAAATQ EVSEAVLALW LPTDSATVLQ KMNEIQAIAR RLPNVDLVLS QTKQDIARAR 960
RLQAEAEER SRRAHVEGVQ EDVVGNLROG TVALQEAQDT MQGTSRSLRL IQDRVAEVQ 1020
VLRPAEKLVT SMTKQLGDFW TRMEELRHQA RQQAEGAVQA QQLAEGASEQ ALSAQEGFER 1080
IKQRYAELKD RLQSSMLGE QGARIQSVKT EAEELFGETM EMDRMKDM ELLLRGSQAI 1140
MLRSADLTGL EKRVEQIRDH INGRVLYYAT CK

Seq ID NO: 21 DNA sequence
Nucleic Acid Accession #: NM_003722
Coding sequence: 145-1491

1 11 21 31 41 51
TCGTGATAT CAAAGACAGT TGAAGGAAAT GAATTTTGAA ACTTCACGGT GTGCCACCCT 60
ACAGTACTCG CCTGACCCCT ACATCCAGCG TTTCTGTAGA ACCCAGCTCA TTTCTCTTGG 120
AAAGAAAGTT ATTACCGATC CACCATGTCC CAGAGCACAC AGACAAATGA ATTCCTCAGT 180
CCAGAGGTTT TCCAGCATAT CTGGGATTTT CTGGAACAGC CTATATGTTC AGTTCAGCCC 240
ATTGACTTGA ACTTTGTGGA TGAACCATCA GAAGATGGTG CGACAAACAA GATTGAGATT 300
AGCATGGACT GTATCGCAT GCAGGACTCG GACCTGAGTG ACCCCATGTG GCCACAGTAC 360
ACGAACCTCG GGCTCCTGAA CAGCATGGAC CAGCAGATTC AGAACGGCTC CTGCTCCACC 420
AGTCCCTATA ACACAGACCA CGCGCAGAAC AGCGTCACGG CGCCCTCGCC CTACGCACAG 480
CCAGCTTCCA CCTTCGATGC TCTCTCTCCA TCACCCGCCA TCCCTCCAA CACCGACTAC 540
CCAGGCCCGC ACAGTTTTCGA CGTGTCTTTC CAGCAGTCCA GCACCGCCAA GTCGGCCACC 600
TGGACGTATT CAGTGAACCT GAAGAACTC TACTGCCAAA TTGCAAGAGC ATGCCCATC 660
CAGATCAAGG TGATGACCCC ACCTCCTCAG GGAGCTGTTA TCCGCGCCAT GCCTGTCTAC 720
AAAAAGCTG AGCAGCTCAC GGAGGTGCTG AAGCGGTGCC CCAACCATGA GCTGAGCGGT 780
GAATTCACAG AGGGACAGAT TGCCCTCTCT AGTCATTGTA TTCGAGTAGA GGGGAACAGC 840
CATGCCCACT ATGTAGAAGA TCCCATCACA GGAAGACAGA GTGTGCTGGT ACCTTATGAG 900
CCAGCCCAAG TTGGCACTGA ATTCAGGACA GTCTTGTACA ATTTCTGTGT TAACAGCAGT 960
TGTGTGGAG GGAATGAACG CCGTCCAATT TTAATCATTG TTAATCTGGA AACCAGAGAT 1020
GGGCAAGTCC TGGGCCGACG CTGCTTTGAG GCCCGGATCT GTGCTTGCCC AGGAAGAGAC 1080
AGGAAGCCCG ATGAAGATAG CATCAGAAAG CAGCAAGTTT CGGACAGTAC AAAGAACCGT 1140
GATGGTACGA AGCGCCCGTT TCGTCAGAAC ACACATGGTA TCCAGATGAC ATCCATCAAG 1200
AAACGAAGAT CCCCAATGA TGAACGTGTA TACTTACCAG TGAGGGGCGG TGAGACTTAT 1260
GAAATGCTGT TGAAGATCAA AGAGTCCCTG GAACTCATGC AGTACCTTCC TCAGACACAA 1320
ATTGAAACGT ACAGGCAACA GCAACAGCAG CAGCACCAGC ACTTACTTCA GAAACATCTC 1380
CTTTCAGCCT GCTTCAGGAA TGAGCTTGTG GAGCCCCGGA GAGAACTCC AAAACAATCT 1440
GAGCTCTTCT TTAGACATTC CAAGCCCCCA AACCGATCAG TGTACCCATA GAGCCCTATC 1500
TCTATATTTT AAGTGTGTGT GTTGTATTTT CATGTGTATA TGTGAGTGTG TGTGTGTGTA 1560
TGTGTGTGCG TGTGTATCTA GCCCTCATAA ACAGGACTTG AAGACACTTT GGCTCAGAGA 1620
CCCACTGTCT CAAAGGCACA AAGCCACTAG TGAGAGAATC TTTTGAAGGG ACTCAAACCT 1680
TTACAAGAAA GGATGTTTTT TGCAGATTTT GTATCCTTAG ACCGGCCATT GGTGGGTGAG 1740
GAACCACTGT GTTGTCTGTG GAGCTTCTGT TTTTTCCTGT GGAGGGAGGG GTCAGGTGGG 1800
GAAAGGGGCA TTAAGATGTT TATTGGAACC CTTTCTGTCT TTTCTCTGTT GTTTTCTTAA 1860
AATTCACAGG GAAGCTTTTG AGCAGTCTC AAACCTAAGA TGTCTTTTAA AGAAAAGGAG 1920
AAAAAGTTG TTATTTGTCTG TGCATAAGTA AGTTGTAGGT GACTGAGAGA CTCACTCAGA 1980
CCCTTTTAAT GCTGGTCAAG TAATAATATT GCAAGTAGTA AGAAACGAAG GTGTCAAGTG 2040
TACTGCTGGC CAGCGAGGTG ATCATTACCA AAAGTAATCA ACTTTGTGGG TGGAGAGTTC 2100
TTTGTGAGAA CTTGCATTAT TTGTGTCTCT CCCTCATGTG TAGGTAGAAC ATTTCTTAAT 2160
GCTGTGTACC TGCCCTGTGC ACTGTATGTT GGCATCTGTT ATGCTAAAGT TTTTCTGTGA 2220
CATGAAACCC TGGAGACCTT ACTACAAAAA AACTGTTGTT TGGCCCCCAT AGCAGGTGAA 2280
CTCATTTTGT GCTTTTAATA GAAAGACAAA TCCACCCGAG TAATATTGCC CTTACGTAGT 2340
TGTTTACCAT TATTCAAAGC TCAAAATAGA ATTTGAAGCC CTCTCAGAAA ATCTGTGATT 2400
AATTGCTCTA ATTAGACTT CTATCCCTCA AGCCTACCTA CCAATAAAAC AGCCATATTA 2460
CTGATACTGT TCACTGCATT TAGCCAGGAG ACTTACGTTT TGAGTAAGTG AGATCCAAGC 2520
AGACGTGTTA AAATCAGCAC TCCTGGACTG GAAATTAAGG ATTGAAAGGG TAGACTACTT 2580

TTCTTTTTT TACTCAAAG TTTAGAGAAT CTCTGTTTCT TTCCATTTTA AAAACATATT 2640
 TTAAGATAAT AGCATAAAGA CTTTAAAAAT GTTCCTCCCC TOCATCTTCC CACACCCAGT 2700
 CACCAGCACT GTATTTTCTG TCACCAAGAC AATGATTCTT TGTATTGAG GCTGTGCTT 2760
 TTGTGGATGT GTGATTTTAA TTTTCAATAA ACTTTTGCAT CTGTGTTTAA AAGAAA

Seq ID NO: 22 Protein sequence:
 Protein Accession #: NP_003713

1 11 21 31 41 51
 MSQSTQTNF LSPEVFQHIW DFLBQPICSV QPIDLNFVDE PSEDGATNKI EISMDCTRMQ 60
 DSDLSDPMNP QYTNLGLLNS MDQQLQNGSS STSPYNTDHA QNSVTAPSPY AQPSTTFDAL 120
 SPSPAPISNT DYPGPHSFDV SFQSSSTAKS ATWTYSTELK KLYCQIAKTC PIQIKVMTTP 180
 PQGAVIRAMP VYKKAHVTE VVKRCPNHEL SREFNEQJIA PPSHLIRVEG NSEAQYVEDP 240
 ITGRQSVLVP YEPQVGTGF TTVLYNFMCI SSCVGGMNRN PILIIVTLET RDGQVLGRRR 300
 FEARICACPG RDRKADEDSE RKQVSDSTK NGDGTKRPFQ QNTHGIQMTS IKKRRSPDDE 360
 LLYLPVRGRE TYEMLLKIKE SLELMQVLPQ HTIETYRQQQ QQQHQHLLQK HLLSACFRNE 420
 LVEPRRETPK QSDVFFRHSK PPNRSVYP

Seq ID NO: 23 DNA sequence
 Nucleic Acid Accession #: NM_001944.1
 Coding sequence: 84-3083

1 11 21 31 41 51
 TTCTCTAGA CATTAAGTGC AGACGGCTGG CAGGATAGAA GCAGCGGCTC ACTTGGAATT 60
 TTTCACACAG GAAATCAGAG ACAATGATGG GGCTCTTCCC CAGAACTACA GGGGCTCTGG 120
 CCATCTTCGT GGTGGTCATA TTGGTTCATG GAGAATTGCG AATAGAGACT AAAGGTCAAT 180
 ATGATGAAGA AGAGATGACT ATGCAACAAG CTAAAAGAAG GCAAAAACGT GAATGGGTGA 240
 AATTGCCCAA ACCCTCAGCA GAAGGAGAAG ATAACTCAAA AAGAAACCCA ATTGCCAAGA 300
 TTACTTCAGA TTACCAAGCA ACCCAGAAAA TCACCTACOG AATCTCTGGA GTGGGAATCG 360
 ATCAGCGGCC TTTTGAATCT TTTGTGTGTG ACAAACACAC TGGAGATATT AACATAACAG 420
 CTATAGTCGA CCGGAGGAGAA ACTCCAAGCT TCCTGATCAC ATGTGCGGCT CTAATGCCCC 480
 AAGGACTAGA TGTAGAGAAA CCACCTTATAC TAACGGTTAA AATTTTGGAT ATTAATGATA 540
 ATCCTCCAGT ATTTTCACAA CAAATTTTCA TGGGTGAAT TGAAGAAAT AGTGCCCTCA 600
 ACTCAGTGTG GATGATACTA AATGCCACAG ATGCAGATGA ACCAAACCCAT TTGAATTTCTA 660
 AAATTGCCCT CAAATTTGTC TCTCAGGAAC CAGCAGGCAC ACCCATGTTT CTCTAAGCA 720
 GAAACACTGG GGAAGTCCGT ACTTTGACCA ATTCTCTTGA CCGAGAGCAA GCTAGCAGCT 780
 ATGCTCTGGT TGTGAGTGGT GCAGACAAAG ATGGAGAAGG ACTATCAACT CAATGTGAAT 840
 GTAATATTAA AGTGAAGAGT GTCAACGATA ACTTCCCAAT GTTTAGAGAC TCTCAGTATT 900
 CAGCAGTAAAT TGAAGAAAAA ATTTTAAAGT CTGAATTACT TCGATTTCAA GTAACAGATT 960
 TGGATGAAGA GTACACAGAT AATTGGCTTG CAGTATATTT CTTTACCTCT GGGAAATGAAG 1020
 GAAATTTGGT TGAATACAAA ACTGATCCTA GAACATAAGA AGGCATCTCT AAAGTGGTGA 1080
 AGGCTCTAGA TTATGAACAA CTACAAAGCG TGAACCTTAG TATTGCTGTC AAAACAAAG 1140
 CTGAATTTCA CCAATCAGTT ATCTCTCGAT ACCGAGTTCA GTCAACCCCA GTCAACATTC 1200
 AGGTAATAAA TGTAAAGAAA GGAATTGCAT TCCGTCCTGC TTCCAAGACA TTACTGTGTC 1260
 AAAAAAGCAT AATGACAAA AAATTGGTGG ATTATATCCT GGGAAACATAT CAAGCCATCG 1320
 ATGAGGACAC TAACAAAGCT GCCTCAATG TCAATATGT CATGGGAGT AACGATGGTG 1380
 GATACCTAAT GATTGATTCA AAAACTGCTG AAATCAAAT TGTCAAAAT ATGAACCGAG 1440
 ATTCTACTTT CATAGTTAAC AAAACATCA CAGCTGAGGT TCTGGCCATA GATGAATACA 1500
 CGGGTAACAA TTCTACAGCA CGGATATATG TTAGAGTACC CGATTTTCAAT GACAATTGTC 1560
 CAACAGCTGT CCTCAAAAAA GATGCAAGTT GCAGTTCTTC ACCTTCCGTG GTTGTCTCCG 1620
 CTAGAACACT GAATAATAGA TACACTGSCC CCTATACATT TGCAGTGGAA GATCAACCTG 1680
 TAAAGTTGCC TGCCGTATGG AGTATCACAA CCCCTCAATG TACCTCGGCC CTCTCAGAG 1740
 CCCAGGAACA GATACCTCCT GGAGTATACC ACATCTCCCT GGTACTTACA GACAGTCAGA 1800
 ACAATCGGTG TGAGATGCCA CGCAGCTTGA CACTGGAAGT CTGTGAGTGT GACAACAGGG 1860
 GCATCTGTGG AACTTCTTAC CCAACACAA GCCCTGGGAC CAGGTATGGC AGGCCGCACT 1920
 CAGGGAGGCT GGGGCTCGCC GCCATCGGCC TGCTGCTCCT TGGTCTCTG CTGCTGCTGT 1980
 TGGCCCCCTT TCTGCTGTTG ACCTGTGACT GTGGGGCAGG TTCTACTGGG GAGATGACAG 2040
 GTGGTTTTAT CCAGTTCCTT GATGGCTCAG AAGGAACAAT TCATCAGTGG GGAATTGAAG 2100
 GAGCCCATCC TGAAGACAA GAAATCACAA ATATTGTGT GCCTCCTGTA ACAGCCAATG 2160
 GAGCCGATTT CATGGAAAGT TCTGAAGTTT GTACAAATAC GTATGCCAGA GGCACAGCGG 2220
 TGAAGGCAC TTCAGGAATG GAAATGACCA CTAAGCTTGG AGCAGCCACT GAATCTGGAG 2280
 GTGCTGCAGG CTTTGCAACA GGGACAGTGT CAGGAGCTGC TTCAGGATTC GAGCAGGCCA 2340
 CTGGAGTTGG CATCTGTTCC TCAGGGCAGT CTGGAACCAT GAGAACAAGG CATTCACATG 2400
 GAGGAACCAA TAAGGACTAC GCTGATGGGG CGATAAGCAT GAATTTTCTG GACTCCTACT 2460
 TTTCTCAGAA AGCATTTGCC TGTGCGGAGG AAGACGATGG CCAGGAAGCA AATGACTGCT 2520
 TGTGTATCTA TGATAATGAA GGCAGAGATG CCACTGTTTC TCTGTGGGC TCCGTGGTT 2580
 GTTGCAAGTT TATTGCTGAT GACCTGGATG ACAGCTTCTT GGACTCACTT GGACCCAAAT 2640
 TTAATAAATC TGACAGATA AGCCTTGGTG TTGATGGTGA AGGCAAAGAA GTTCAGCCAC 2700
 CCTCTAAGA CAGCGGTTAT GGGATTGAAT CCTGTGSCCA TCCCATAGAA GTCCAGCAGA 2760
 CAGGATTTGT TAAGTCCAG ACTTTGTGAG GAAGTCAAGG AGCTTCTGCT TTGTCCGCT 2820
 CTGGGTCTGT CCAGCCAGCT GTTTCATCC CTGACCTCTG GCAGCATGCT AACTATTAG 2880
 TAAAGGAGAC TTAAGTGGCT TCTGGTTCCT TCGTGAACCT TTCACTGCA GGCCTTGTATC 2940
 CACTTCTCAC AAAAAATGTG ATAGTGACAG AAAGGGTGAT CTGTCCCATC TCCAGTGTTC 3000
 CTGGCAACCT AGCTGGCCCA CGCAGCTAC GAGGGTCACA TACTATGCTT TGTACAGAGG 3060
 ATCCTGTGCT CCGTCTAATA TGACCAGAAT GAGCTGGAAT ACCACACTGA CCAATCTGG 3120
 ATCTTTGAGC TAAAGTATTC AAAATAGCAT AGCAAAGCTC ACTGTATTGG GCTAATAATT 3180
 TGGCACTTAT TAGCTTCTCT CATAAATCGA TCAGGATTAT AAATTAAATG TTTGGGTTCA 3240
 TACCCCAAAA GCAATATGTT GTCACTCCTA ATTTCTCAAGT ACTATTCAA TTTAGTAA 3300
 TCTTAAAGTT TTCAAACACC CTAAATCAT ATTGCG

Seq ID NO: 24 Protein sequence:
 Protein Accession #: NP_001935.1

1 11 21 31 41 51

	MGLFPRTTG	ALAIFFVVIL	VHGELEIETK	GOYDEEBMTM	QOAKRRQKRE	WVKFAKPCRE	60
	GEDNSKRNPI	AKITSDYQAT	QKITYSRISGV	GIDQPPFGIF	VVDKNTGDIIN	ITAIVDREET	120
5	PSFLITCRAL	NAQGLDVEKP	LILTVKILDI	NDNPPVFSQQ	IFMGEIEBENS	ASNSLVMILN	180
	ATDADEPNHL	NSKIAFKIVS	QEPAGTMMFL	LSRNTGEVTR	LNSLDREQA	SSYRLVVSQA	240
	DKDGEGLSTQ	CECNIKVKVD	NDNPFMPFRDS	QYSARIEENI	LSSELLRFQV	TDLDEEYTDN	300
	WLAVYFFTSQ	NEGNWFEIQT	DPRINEGILK	VVKALDYEQL	QSVKLSIAVK	NKAEFHQSVI	360
	SRYRVQSTPV	TIQVINVREG	IAPRPASKTF	TVQKGISKK	LVDYILGTQY	AIDEDTHKAA	420
10	SNVYVVMGRN	DGGYLMIDSK	TAEIKFVKNM	NRDSTFIYNK	TIITAEVLAI	EYTGKTSTGT	480
	VYVRVPDFND	NCPTAVLEKD	AVCSSSPSVV	VSARTLNNRY	TGPTTFALD	QPVKLPVWS	540
	ITTLNATSAL	LRAQEQIPPG	VYHISLVLT	SQNNRCENPR	SLTLEVCQCD	NRGICGTSYP	600
	TTSPTGTRYGR	PHSGRLGPA	IGLLLLGLLL	LLLAFLILLT	CDCGAGSTGG	VTGGFIFVPD	660
	GSEGTIHQWG	IEGAHPEDKE	ITNICVPPVT	ANGADFMESS	EVCTNTYARG	TAVEGTSME	720
15	MTTKLGAATE	SGGAAGFATG	TVSGAASGFG	AATGVGICSS	QSGGTMTRTH	STGGTNKDYA	780
	DGAISMNPLD	SYFSQKAFAC	AEEDDQGEAN	DCLLIYDNEG	ADATGSPVGS	VGCCSFIADD	840
	LDDSFLLSLG	PKFKKLAEIS	LGVDGEGKEV	QPPSKDSGYG	IESQGHPIEV	QQTGFVKCQT	900
	LSGSGGASAL	SASGSVQPAV	SIPDPLQHG	YLVETETYSAS	GSLVQPSLAG	FDPLLTQNV	960
	VTERRVICPIS	SVFNGLAGPT	QLRGSHMTMLC	TEDPCSRLI			

20 Seq ID NO: 25 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 56-1642

	1	11	21	31	41	51	
25	AGTATCCAG	GAGGAGCAAG	TGGCAGCTCT	TCGGACCTAG	GCTGCCCTG	CCGTATGTC	60
	GCAAGGGATC	CTTCTCCG	CAGCGGGCTT	GCTGTCCGAT	GACGATGTC	TAGTTCTCC	120
	CATGTTTGAG	TCCACAGCTG	CAGATTGGGG	GTCTGTGGTA	GCAGAGAAC	TGCTATCAGA	180
	CTGCTCTGTC	GTCTCTACCT	CCCTAGAGGA	CAAGCAGCAG	GTTCCATCTG	AGGACAGTAT	240
30	GGAGAAGGTG	AAAGTATAT	TGAGGGTTAG	GCCCTTGTTA	CCTTCAGAGT	TGGAACGACA	300
	GGAAGATCAG	GTTGTGTGCC	GTATTGAGAA	TGTGGAGACC	CTTGTCTAC	AAGCACCCAA	360
	GGACTCTTTT	GCCCTGAAGA	GCAATGAACG	GGGAATTGGC	CAAGCCACAC	ACAGGTTTAC	420
	CTTTTCCAG	ATCTTTGGGC	CAGAAGTGGG	ACAGGCATCC	TCTTCAACC	TAAGTGTGAA	480
	GGAGATGGTA	AAGGATGTAC	TCAAAGGGCA	GAATCTGGCT	ATCTATACAT	ATGGAGTAC	540
35	TAATCTAGGG	AAAACCCACA	CGATTCAAGG	TACCATCAAG	GATGGAGGGA	TTCTCCCGCG	600
	GTCCTCTGGC	CTGATCTTCA	ATAGCCTCCA	AGGCCAACTT	CATCCACAC	CTGATCTGAA	660
	GCCCTTGCTC	TCCATAGAGG	TAATCTGGCT	AGACAGCAAG	CAGATCCGAC	AGGAGGAAAT	720
	GAAGAAGCTG	TCCCTGCTAA	ATGGAGGCTT	CCAAGAGGAG	GAGCTGTCCA	CTTCTTGAA	780
40	GAGGAGTGTC	TACATCGAAA	GTCCGATAGG	TACCAGCACC	AGCTTCGACA	GTGGCATTGC	840
	TGGGCTCTCT	CTATCTAGTC	AGTGTACCAG	CAGTAGCCAG	CTGATGAAA	CAAGTCTATG	900
	ATGGGCACAG	CCAGACACTG	CCCCACTACC	TGTCCCGGCA	AACATTGCTT	TCTCCATCTG	960
	GATCTCATTC	TTTGAGATCT	ACAACGAACT	GCTTTATGAC	CTATTAGAAC	CGCCTAGCCA	1020
	ACAGCGCAAG	AGGCAGACTT	TGCGGCTATG	CAGGAGTCAA	AATGGCAATC	CCTATGTGAA	1080
45	AGATCTCAAC	TGGATTTCAT	TGCAAGATGC	TGAGGAGGCC	TGGAAGCTCC	TAAAGTGGG	1140
	TGTAAGAAGC	CAGAGCTTTC	CCAGCACCCA	CCTCAACCCG	AACCTCCAGC	GCAGTCAACG	1200
	CATCTTCTCA	ATCAGGATCC	TACACCTTCA	GGGGGAAGGA	GATATAGTCC	CCAAGATCAG	1260
	CGAGCTGTCA	CTCTGTGATC	TGGCTGGCTC	AGAGCGCTGC	AAAGATCAGA	AGAGTGGTGA	1320
	ACGGTTGAAG	GAAGCAGGAA	ACATTAAAC	CTCTCTACAC	ACCCCTGGCC	GCTGTATTGC	1380
50	TGCCCTTCGT	CAAAACCCAG	AGAACCGGTC	AAAGCAGAAC	CTGGTTCCCT	TCCGTGACAG	1440
	CAAGTTGACT	CAGGTGTTC	AAGGTTTCTT	CACAGGCCGA	GGCCGTTCTT	GCATGATTGT	1500
	CAATGTGAAT	CCCTGTGCAT	CTACCTATGA	TGAAACTCTT	CATGTGGCCA	AGTTCTCAGC	1560
	CATTGTGAGC	CAGGTGACTT	GTGCATGCC	CACCTATGCA	ACTGGGATTC	CCATCCCTGC	1620
	ACTCGTTTAT	CAAGGAACAT	AGTCTTTCAG	TATCCCCACG	CTTAGAGAAA	GGGGCTAAGG	1680
55	CAGACACAGG	CCTTGATGAT	GATATTGAAA	ATGAAGCTGA	CATCTCCATG	TATGGCAAGG	1740
	AGGAGCTCCT	CAAGTTTGTG	GAAGCCATGA	AGACACTGCT	TTTGAAGGAA	CGACAGGAAA	1800
	AGCTACAGCT	GGAGATGCAT	CTCCGAGATG	AAATTTGCAA	TGAGATGGTA	GAACAGATGC	1860
	AACAGCGGGA	ACAGTGGTGC	AGTGAACATT	TGACACCCA	AAAGGAACCT	TTGGAGGAAA	1920
	TGTAAGAAGA	AAAACATAAT	ATCTCAAGG	AGTCACTGAC	AAGTTTTTAC	CAAGAAGAGA	1980
60	TTCAAGGAGC	GGATGAAAGC	ATTGAAGAGC	TAGAAGCTCT	CTTGACAGAA	GCCAGACAAC	2040
	AGTCAGTGGC	CCATCAGCAA	TCAGGGTCTG	AATTGGCCCT	ACGGCGGTCA	CAAGAGTTGT	2100
	CAGCTTCTGC	CTCCACCCAG	CAGCTTCAGG	AGGTTAAAGC	TAAATTACAG	CAGTGCAGAG	2160
	CAGAGCTAAA	CTCTACCACT	GAAGAGTTGC	ATAAGTATCA	GAAATGTGTA	GAACCAACAC	2220
	CCTCAGCCAA	GCCCTTCACC	ATTGATGTGG	ACAAGAAGTT	AGAAGAGGGC	CAGAAGAATA	2280
65	TAAGGCTGTT	GCGGACAGAG	CTTCAGAAAC	TTGGTGAGTC	TCTCCAATCA	GCAGAGAGAG	2340
	CTTGTGGCCA	CAGCACTGGG	GCAGGAAAAC	TTCTGCAAGC	CTTGACCACT	TGTGATGACA	2400
	TCTTAATCAA	ACAGGACCCG	ACTCTGGCTG	AACTGCAGAA	CAACATGGTG	CTAGTGAAAC	2460
	TGGACCTTCG	GAAGAAGGCA	GCATGTATTG	CTGAGCAGTA	TCATCTGTG	TTGAAACTCC	2520
	AAGGCCAGGT	TTCTGCCAAA	AAGCGCCTTG	GTACCAACCA	GGAAAATCAG	CAACCAAAAC	2580
70	AACAACCAAC	AGGGAAGAAA	CCATTCTCTC	GAAATTTACT	TCCCGGAAAC	CCAACCTGCC	2640
	AAAGCTCAAC	AGACTGCAGC	CCTTATGCC	GGATCCTACG	CTCACGGCGT	TCCCTTTTAC	2700
	TCAAATCTGG	GCCCTTTTGG	AAAAAGTACT	AAGGCTGTGG	GGAAAGAGAA	GAGCAGTCAT	2760
	GGCCCTGAGG	TGGGTACAGT	ACTCTCTCTG	AGAAATAGGT	CTCTTTTATG	CTTTACCAT	2820
	TATCAGGAAT	TATATCCAGG	ATGCAATACT	CAGACACTAG	CTTTTCTCTC	ACTTTTGTAT	2880
75	TATAACCAAC	TATGTAATCT	CATGTTGTTG	TTTTTTTTTA	TTTACTTATA	TGATTTCTAT	2940
	GCACACAAA	ACAGTTATAT	TAAAGATATT	ATTGTTTACA	TTTTTTATTG	AATTCCAAAT	3000
	GTAGCAAAAT	CATTAATAA	AATTATAAAA	GGGACAGAAA	AA		

80 Seq ID NO: 26 Protein sequence:
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
85	MSQGLSPPPA	GLLSDDDVVV	SPMFESTAAD	LGSVVRKNLL	SDCSVVSTSL	EDKQVPSSED	60
	SMEKVYVYLR	VRPLLPSLEL	RQEDQGCVR	ENVETLVLAQ	PKDSFALKSN	ERGIGQATHR	120
	PTFSQIFQPE	VQASFPNLTV	VKEMVKDVLK	GQNLVIYTYG	VNSGKTHTI	OGTIKDGGLL	180
	PRSLALIFNS	LQGLHPTPD	LKPLLSNEVI	WLDKQIRQ	EMKRLSLNLG	GLQEELSTLS	240
	LKRSVYIESR	IGTSTSPDSG	IAGLSSISQC	TSSSQLDETS	HRWAQPDTPA	LPVPANIRFS	300

IWISFFBIYN ELLYDLEPP SQQRKROTLR LCEDQNGNPF VKDLNWIHVO DAEBAWKLLK 360
VGRKNQSPAS THLNQNSRRS HSIIFSIRILH LQEGSDIVPK ISELSLCDLA GSERCKDQKS 420
GERLKEAGNI NTSLHTLGRG IAAIRLQNOQN RSKQNLVPPR DSKLITRVFQG FFTGRGRSCM 480
IVNVNPCAST YDETLHVAKF SAIASQVTC APTYATGIPI PALVHQGT

Seq ID NO: 27 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 13-1424

10 1 11 21 31 41 51
TAGAAGTTTA CAATGAAGTT TCTTCTAATA CTGCTCCTGC AGGCCACTGC TTCTGGAGCT 60
CTTCCCTCGA ACAGCTCTAC AAGCCTGGAA AAAAATAATG TGCTATTGG TGAAAGATAC 120
TTAGAAAAAT TTTATGGCCT TGAGATAAAC AAATCTCCAG TGACAAAAAT GAAATATAGT 180
15 GGAAACTTAA TGAAGGAAAA AATCCAAGAA ATGCAGCACT TCTGGGTCT GAAAGTGACC 240
GGGCACCTGG ACACATCTAC CCTGAGATG ATGCAGCAC CTGATGTGG AGTCCCGAT 300
GTCCATCAT TCGAGGAAAT GCCAGGGGGG CCGGTATGGA GAAACATTA TATCACCTAC 360
AGAATCAATA ATTACACACC TGACATGAAC CGTGAGGATG TTGACTACGC AATCCGGAAA 420
GCTTTCACAT TATGAGTAA TGTATACCCC TTGAAATTC ACAAATTAAC CACAGGCATG 480
20 GCTGACATTT TGGTGGTTTT TGCCCGTGGG GCTCATGGAG ACTTCCATGC TTTTGTATGG 540
AAAGGTGGAA TCCTAGCCCA TGTCTTTTGA CTGCTATCTG GCATTGGAGG GATGTCACAT 600
TTGATGAGG ACGAATTTCT GACTACACAT TCGAGAGGCA CAACTTGTG CCTCACTGCT 660
GTTACAGAGA TTGCCCATCT CTTAGGTCCT GGCCATTCTA GTGATCCAAA GGCGGTAAAT 720
TTCCCACTCT ACAATATGT TGACATCAAC ACATTTGGCC TCTCTGCTGA TGACATACGT 780
25 GGCATTCACT CCTGTATGT AGACCCAAAA GAGAACCAAC GCTTGCCAAA TCCTGACAA 840
TCAGAACCG CTCTCTGTGA CCCCAATTG AGTTTGTATG CTGTCACTAC CGTGGGAAAT 900
AAGATCTTTT TCTTCAAGA CAGGTTCTTC TGGCTGAAGG TTTCTGAGAG ACCAAAGACC 960
AGTGTTAATT TAATTTCTTC CTTATGGCCA ACCTTGCCAT CTGGCATGTA AGCTGCTTAT 1020
30 GAAATGAAG CCAGAAATCA AGTTTCTCT TTTAAAGATG ACAAATACTG GTTAATTAGC 1080
AATTTAAGC CAGAGCCAAA TTATCCCAAG AGCATACATT CTTTGGTGT TCCTAACTTT 1140
GTGAAAAAAA TTGATGCAGC TGTTTTAAAC CCACGTTTT ATAGGACCTA CTTCTTTGTA 1200
GATAACCACT ATTTGGAGTTA TGATGAAAGG AGACAGATGA TGGACCTCGG TTATCCCAAA 1260
CTGATACCA AGAACTTCCA AGGAATCGGG CCTAAAATG ATGCAGTCTT CTACTCTAAA 1320
35 AACAAATACT ACTATTTCT CCAAGGATCT AACCAATTG AATATGACTT CTACTCTAAA 1380
CGTATACCA AACCACTGAA AAGCAATAGC TGGTTTGGTT GTTGAAATG GTGTAATTA 1440
TGGTTTTTGT TAGTTCACTT CAGCTTAATA AGTATTTATT GCATATTGTC TATGCTCTCA 1500
GTGTACCACT ACTTAGAGAT ATGTATCATA AAAATAAAAT CTGTAAACCA TAGGTAATGA 1560
TTATATAAAA TACATAATAT TTTTCAATT TGAAACTCT AATTGTCCAT TCTTGCTTGA 1620
CTCTACTATT AAGTTTGAAA ATAGTTACCT TCAAGCAAG ATAATTCTAT TGAAGCATG 1680
40 CTCGTAAAGT TGTCTCTAAA CATCTTGA CTGAGAAAT ATACTTACTT CTGCGATAAC 1740
TAAATTAAG TATATATATT TTGGCTCAAA TAAATTTG

Seq ID NO: 28 Protein sequence:
Protein Accession #: Eos sequence

45 1 11 21 31 41 51
MKFLILLIQ ATASGALPLN SSTSLKNNV LFGERYLEK YGLEINKLPV TKMKYSGNLM 60
KEKIQEMQRF LGKLVGTQGLD TSTLEMMHAP RCGVPDVHVF REMPGGPVWR KHYITYRIIN 120
50 YTPDMREOV DYAIRKAFQV WSNVPLPKFS KINTGMADIL VVFARGAHGD PHAFDGGKGI 180
LAHAFGPGSG IGGDAHFEDD EFWTTESGGT NLFITAVHEI GHSILGLHSS DPKAVMFPTY 240
KYVDINTFRL SADDIRGIQS LYGDPEKQNR LPNPDNSEPA LCDPNLSFDA VTTVGNKIFF 300
FKDRPFWLVK SERPKTSVNL ISSLWPTLPS GIEAAYEIEA RNQVFLPKDD KYWLISNLRP 360
EPNPKSIHS FGPFNVVKKI DAAVFNPRFY RYTFVVDNQY WRDERRQMM DPGYFKLITK 420
55 NFQIGIPKID AVFYSKNKYY YFPQGSNQFE YDFLLQRIK TLKSNWFHGC

Seq ID NO: 29 DNA sequence
Nucleic Acid Accession #: NM_006115.1
Coding sequence: 236..1765

60 1 11 21 31 41 51
GCTTCAGGGT ACAGCTCCCC CGCAGCCAGA AGCCGGGCGT GCAGCCCTCT AGCACCGCTC 60
CGGACACACC CACCCGCTTC CCAGGCGTGA CCTGTCAACA GCAACTTCGC GGTGTGGTGA 120
65 ACTCTCTGAG GAAAAACCAT TTTGATTATT ACTCTCAGAC GTGCGTGGCA ACAAGTGACT 180
GAGACCTAGA AATCCAAGCG TTGGAGGTCC TGAGGCCAGC TAAAGTCTGT TCAAAATGGA 240
ACGAAGCGGT TTGTGGGGTT CCATTGAGAG CCGATACATC AGCATGAGTG TGTGGACAAG 300
CCCAAGGAGA CTTGTGGAGC TGGCAGGGCA GAGCCTGCTG AAGGATGAGG CCCTGGCCAT 360
TGCGCCCTCG GAGTTGCTGC CCAGGGAGCT CTTCCGCCCA CTCTTCATGG CAGCCTTTGA 420
70 CGGAGACAC AGCCAGACCC TGAAGGCAAT GGTGCAGGCC TGGCCCTTCA CTTGCTCTCC 480
TCTGGGAGTG CTGATGAAGG GACAACATCT TCACCTGGAG ACCTTCAAAG CTGTGCTTGA 540
TGGACTTGAT GTGCTCTCTG CCCAGGAGGT TCGCCCGCAG AGGTGGAAC TTCAAGTGCT 600
GGATTTACGG AAGAACTCTC ATCAGGACTT CTGGACTGTA TGGTCTGGAA ACAGGGCCAG 660
75 TCTGTACTCA TTTCCAGAGC CAGAAGCAGC TCAGCCCATG ACAAAGAAGC GAAAGTAGA 720
TGGTTTGAGC ACAGAGGCGA AGCAGCCCTT CATTCCAGTA GAGTGCTCG TAGACCTGTT 780
CCTCAAGGAA GGTGCTGTG ATGAATTGTT CTCTCACTCT ATTGAGAAAG TGAAGCGAAA 840
GAAAAATGTA CTAGCCTGT GCTGTAAGAA GCTGAAGATT TTTGCAATGC CCATGCAAGA 900
TATCAAGATG ATCTGAAAG TGGTGCAGCT GGACTCTATT GAAGATTGG AAGTGACTTG 960
TACCTGGAAG CTACCCACCT TGGCGAAAT TTTCTCTTAC CTGGGCCAGA TGATTAACT 1020
80 GCGTAGACTC CTCTCTCCC ACATCCATGC ATCTTCTTAC ATTTCCCGG AGAAGGAAGA 1080
GCAGTATATC GCCCAGTTCA CCTCTCAGTT CCTCAGTCTG CAGTGCTCTG AGGCTCTCTA 1140
TGTGACTCT TATTTTCTCC TTAGAGCCCG CCTGGATCAG TTGCTCAGGC ACGTGATGAA 1200
CCCTTTGAAA ACCCTCTCAA TAACTAACTG CCGGCTTTCG GAAGGGGATG TGATGCACT 1260
85 GTCCAGAGT CCCAGGCTCA GTCAGCTAAG TGTCTGTAGT CTAAGTGGGG TCAATGCTAG 1320
CGATGTAAGT CCGAGCCCC TCCAAGCTCT TGATCAGCTC CTGCGCTTCC TGCCTTCCCT 1440
CCTGGTCTTT GATGAGTGTG GGATCAGGGA GATCAGCTC CTGCGCTTCC TGCCTTCCCT 1500
GAGCCACTGC TCCACAGTTA CAACCTTAAG CTTCTAGGG AATTCCATCT CCATATCTGC

5 CTTCAGAGT CTCTGCAGC ACCTCATCGG GCTGAGCAAT CTGACCCAGC TGCTGTATCC 1560
 TGTCCCTCTG GAGAGTTATG AGGACATCCA TGGTACCTTC CACTGGAGA GGCTTGCTTA 1620
 TCTGCATGCC AGGCTCAGGG AGTTGCTGTG TGAGTTGGGG CGGCCAGCA TGGTCTGGCT 1680
 TAGTGCCAAC CCCTGTCTTC ACTGTGGGGA CAGAACTTC TATGACCCGG AGCCCATCCT 1740
 GTGCCCTGT TTCTAGCCTA ACTAGCTGGG TGCACATATC AAATGCTTCA TTCTGCATAC 1800
 TTGGACACTA AAGCCAGGAT GTGCATGCAT CTGAAGCAA CAAGCAGCC ACAGTTTCAG 1860
 ACAATGTTC AGTGTGAGTG AGGAAAACAT GTTCAGTGAG GAAAAACAT TCAGACAAAT 1920
 GTTCAGTGAG GAAAAAAGG GGAAGTTGGG GATAGGCAGA TGTGTACTTG AGGAGTTAAT 1980
 10 GTGATCTTTG GGGAGATACA TCTTATAGAG TTAGAAATAG AATCTGAATT TCTAAAGGGA 2040
 GATTCTGGCT TGGGAAGTAC ATGTAGGAGT TAATCCCTGT GTAGACTGTT GTAAAGAAAC 2100
 TGTGAAAT AAAGAGAAGC AATGTGAAGC AAAAAA AAAA

Seq ID NO: 30 Protein sequence:
 Protein Accession #: NP_006106.1

15 1 11 21 31 41 51
 GCTTCAGGTT ACAGCTCCCC CGCAGCCAGA AGCCGGGCGT GCAGCGCCTC AGCACCGCTC 60
 CGGGACAGCC CACCCGCTTC CCAGGCGTGA CCTGTCAACA GCAACTTCGC GGTGTGGTGA 120
 20 ACTCTCTGAG GAAAAACCAT TTTGATTATT ACTCTCAGAC GTGCGTGGCA ACAAGTGACT 180
 GAGACCTAGA AATCCAGCG TTGAGGTCC TGAGGCCAGC CTAAGTCGT TCAAAATGGA 240
 ACGAAGCGGT TTGTGGGTTT CCATTAGAG CCGATACATC AGCATGAGTG TGTGGACAAG 300
 CCCACGGAGA CTTGTGGAGC TGGCAGGGCA GAGCCTGTCT AAGGATGAGG CCTTGGCCAT 360
 25 TGCCGCGCTG GAGTTGCTGC CAGGGAGCT CTCTCCGCCA CTCTCATGG CAGCCTTTGA 420
 CGGGAGACAC AGCCAGACCC TGAAGCAAT GGTGAGGCC TGGCCCTTCA CCTGCTCCC 480
 TCTGGGAGTG CTGATGAAGG GACAACATCT TCACCTGGAG ACCTTCAAAG CTGTCTTGA 540
 TGGACTGTAT GTGCTCCTTG CCCAGGAGT TCGCCCGCAG AGGTGGAAC TTCAAGTGTCT 600
 GGATTTACGG AAGAATCTCT ATCAGGACT CTGGACTGTA TGGTCTGGAA ACAGGGCCAG 660
 TCTGTACTCA TTTCCAGAG CAGAAGCAGC TCAGCCCATG ACAAGAGAGC GAAAAGTAGA 720
 30 TGGTTTGAGC ACAGAGGAGC AGCAGCCCTT CATTCAGTA GAGGTGCTCG TAGACCTGTT 780
 CCTCAAGGAA GGTGCTCTGT ATGAATTGTT CTCTACCTC ATTGAGAAAG TGAAGCGAAA 840
 GAAAAATGTA CTACGCTGT GCTGTAAGAA GCTGAAGATT TTTGCAATGC CCATGCAGGA 900
 TATCAAGATG ATCTCTGAAA TGGTGAGCT GGACTCTATT GAAGATTGG AAGTGACTTG 960
 35 TACCTGGAAG CTACCCACCT TGGCGAAATT TTCTCCTTAC CTGGGCCAGA TGATTAACTCT 1020
 GCGTAGACTC CTCTCTCCC ACATCCATGC ATCTCTCTAC ATTTCCCGG AGAAGGAAGA 1080
 GCAGTATATC GCCAGTTTCA CCTCTCAGTT CCTCAGTCTG CAGTGCCTGC AGGCTCTCTA 1140
 TGTGGACTCT TTATTTTCCC TTAGAGGCCG CCTGGATCAG TTGCTCAGGC ACGTGATGAA 1200
 40 CCCCTTGGAA ACCCTCTCAA TAACCTAAGT CCGGCTTTCC GAAGGGGATG TGATGCATCT 1260
 GTCCCAAGAT CCCAGCTCA GTCACTAAG TGTCTGAGT CTAAGTGGGG TCATGCTGAC 1320
 CGATGTAAAT CCGAGGCCCC TCCAAGCTCT GCTGGAGAGA GCCTCTGCCA CCTCCAGGA 1380
 CCTGGTCTTT GATGAGTGTG GGATCACGGA TGATCAGCTC CTGCGCTCC TGCTTCCCT 1440
 GAGCCACTGC TCCAGCTTAA CAACCTTAAG CTCTACGGG AATTCCATCT CCATATCTGC 1500
 45 CTTGAGAGT CTCTCAGC ACCTCATCGG GCTGAGCAAT CTGACCCAGC TGCTGTATCC 1560
 TGTCCCTCTG GAGAGTTATG AGGACATCCA TGGTACCTTC CACTGGAGA GGCTTGCTTA 1620
 TCTGCTGCC AGGCTCAGG AGTTGCTGTG TGAGTTGGGG CGGCCAGCA TGGTCTGGCT 1680
 TAGTGCCCA CCCTGTCTCT ACTGTGGGGA CAGAACTTC TATGACCCGG AGCCCATCCT 1740
 GTGCCCTGT TTATGCTCTA ACTAGCTGGG TGCACATATC AAATGCTTCA TTCTGCATAC 1800
 50 TTGGACATA AAGCAGGAT GTGCATGCAT CTTGAAGCAA CAAGCAGCC ACAGTTTCAG 1860
 ACAATGTTC AGTGTGAGTG AGGAAAACAT GTTCAGTGAG GAAAAACAT TCAGACAAAT 1920
 GTTCAGTGAG GAAAAAAGG GGAAGTTGGG GATAGGCAGA TGTGTACTTG AGGAGTTAAT 1980
 GTGATCTTTG GGGAGATACA TCTTATAGAG TTAGAAATAG AATCTGAATT TCTAAAGGGA 2040
 GATTCTGGCT TGGGAAGTAC ATGTAGGAGT TAATCCCTGT GTAGACTGTT GTAAAGAAAC 2100
 TGTGAAAT AAAGAGAAGC AATGTGAAGC AAAAAA AAAA

Seq ID NO: 31 DNA sequence
 Nucleic Acid Accession #: E08 sequence
 Coding sequence: 64-2754

60 1 11 21 31 41 51
 GGCAGGTCTC GCTCTGGGCA CCCTCCCGGC GCCCGGTTT TCCTGGCCCT GCCCGGCATC 60
 CCGATGGCCG CCGCTGGGCC CCGGCGCTCC GTGCGGGAG CCGTCTGCTC GCATCTGCTG 120
 65 CTGACCCCTG TGATCTTCAG TCGTATGTT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180
 CTTCTAAAC TAGAGGACAG CAAAATAATT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240
 TCTGCAGACC TCATCCGCTC AAGTGATCCT GATTTCAGAG TTCTAAATGA TGGGTCAAGT 300
 TACACAGCCA GGGCTGTGTC GCTGTCTGAT AAGAAAAGAT CATTTACCAT ATGGCTTTCT 360
 GACAAAAGGA AACAGACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420
 70 TCGAAGACAA GACACACTAG AGAAACTGTT CTCAGGCGTG CCAAGAGGAG ATGGGCACCT 480
 ATTCTTGCT CTATGCAAGA GAATTCCTTG GGCCTTTTCC CATTTTCT TCAACAAGTT 540
 GAATCTGATG CAGCAGACAA CTATCTGTC TTCTACTCAA TAAGTGAGC TGGAGTTGAT 600
 AAAGAACCTT TAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTATT TTGCACTCGG 660
 CCTGTGGATC GTGAAGAATA TGATGTTTTT GATTGTATG CTTATGCGTC AACTGCAGAT 720
 75 GGATATTGAG CAGATCTGCC CCTCCCACTA CCCATCAGGG TAGAGGATGA AAATGACATC 780
 CACCCGTGTT TCACAGAAGC AATTTATAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840
 ACTACAGTGG GGGTGGTTTG TGCCACAGAC AGAGATGAAC CGGACACAT GCATACGCGC 900
 CTGAAATACA GCATTTTGCA GCAGACACCA AGGTCACTCT GGCTCTTTTC TGTGCATCCC 960
 AGCAGGCGG TAATCACCAC AGTCTCTCAT TATTGGACA GAGAGGTTGT AGACAAGTAC 1020
 80 TCATTGATAA TGAAGTACA AGACATGGAT GGCAGTTT TTGGATTGAT AGGCACATCA 1080
 ACTTGTATCA TAACAGTAAC AGATTCAAAT GATAATGCAC CCACTTTCAG ACAAATGCT 1140
 TATGAAGCAT TTGTAGAGGA AATGTCATT AATGTGAAA TCTTACGAAT ACCTATAGAA 1200
 GATAAGGATT TAATTAAAC TGCCAATTG AGAGTCAATT TTACCATTTT AAAGGGAAT 1260
 GAAAATGGAC ATTTTCAAAT CAGCAGACAC AAGAAAATA ATGAAGTGT TCTTCTGTT 1320
 85 GTAAAGCCAC TGAATTATGA AGAAACCGT CAAGTGAACC TGGAAATGG AGTAAACAT 1380
 GAAGCGCCAT TTGCTAGAGA TATCCCAAGA GTGACAGCCT TGAACAGAGC CTGGTTACA 1440
 GTTCATGTA GGGATCTGGA TGAGGGGCTT GAATGCATCT CTGACGCCCA ATATGTGCGG 1500
 ATTAAGAAA ACTTAGCAGT GGGGTCAAAG ATCAACGGCT ATAAGGCATA TGACCCGAA 1560

	AATAGAAATG	GCAATGGT	TTT	AAGGTACAAA	AAATTGCATG	ATCCTAAAGG	TTGGATCACC	1620
	ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAATCC	TGGATAGGGA	GGTTGAAACT	1680	
	CCCAAAAATG	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740	
5	ACTGGAAACAC	TTGCTGTGAA	CATTGAAAGAT	GTAATGATA	ATCCACCAGA	AATACTTCAA	1800	
	GAATATGTAG	TCATTGGCAA	ACCAAAAATG	GGGTATACCG	ACATTTIAGC	TGTTGATCCT	1860	
	GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTGAGTTTGC	CCAATACTTC	TCCAGAAATC	1920	
	AGTAGACTGT	GGAGCCTCAC	CAAAGTTAAT	GATACAGCTG	CCGCTCTTTC	ATATCAGAAA	1980	
	AATGCTGGAT	TTCAAGAATA	TACCATTCTC	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040	
10	GCAACAAAAT	TATTGAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACTCA	GTGTGCTGGG	2100	
	ACTTCAAGGA	GTACAGGAGT	AATACTTGGA	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160	
	ATAGCACTGC	TCCTTTCTGT	ATTGCTAACT	TTAGTATGTG	GAGTTTTTGG	TGCAACTAAA	2220	
	GGGAAACGTT	TTCTGGAAGA	TTTAGCAGAG	CAAAACTTAA	TTATATCAAA	CACAGAAGCA	2280	
	CCTGGAGACG	ATAGAGTGTG	CTCTGCCAAT	GGATTATGTA	CCCAAACTAC	CAACAACTCT	2340	
15	AGCCAGAGTT	TTTGTGGTAC	TATGGGATCA	GGAAATGAAA	ATGGAGGGCA	GGAAACCAT	2400	
	GAAATGTGA	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCC	GGGGGGCTGG	GCATCATCAT	2460	
	ACCTGGAGCT	CCTGAGGGGG	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACCTACTCG	2520	
	GAGTGGCACA	GTTTTACTCA	ACCCCGTCTC	GGTGA AAAAT	TGCATGATG	TAATCAGAA	2580	
	GAAGACCCGA	TGCCATCCCA	AGATTATGTC	CTCACTTATA	ACTATGAGGG	AAGAGGATCT	2640	
20	CCAGCTGGTT	CTGTGGGCTG	CTGCAGTGAA	AAGCAGGAAG	AAGATGGCCT	TGACTTTTAA	2700	
	AATAAATTTG	AACTCAAGAT	TATTACATTA	GCAGAAAGCAT	GCACAAAGAG	ATAATGTGAC	2760	
	AGTGTCTACA	TTAGTCTTTT	GTGAGACATT	CTGGAGGTTT	CCAAAAATAA	TATTGTAAAG	2820	
	TTCAATTTC	ACATGTATGT	ATATGATGAT	TTTTTCTCA	ATTTTGAATT	ATGCTACTCA	2880	
	CCAAATTTATA	TTTTTAAAGC	CAGTTGTGTC	TTATCTTTTC	CAAAAAGTGA	AAAAATGTTAA	2940	
25	AACAGACAAAC	TGGTAAATCT	CAAACTCCAG	CACCTGGAAT	AAGTCTCTTA	AAGCATCTGC	3000	
	TCCTTTTTTT	TTTACGGGAT	ATTTTAGTAA	TAAATATGCT	GGATAAATAT	TAGTCCAAACA	3060	
	ATAGCTAAGT	TATGCTAATA	TCACATTATT	ATGATTTCAC	TTTAAAGTAT	AGTTTAAAAA	3120	
	ATAAACAGAA	AATATTGAGT	ATCATTATGT	GAAGAAAGTT	TTGGA AAAA	AACAAATGAAG	3180	
	ACTGAATTTAA	ATTAAAAAATG	TTGCAGCTCA	TAAAGAAATG	GGACTCACCC	CTACTGCAC	3240	
30	ACCAAAATCA	TTTGACTTTG	GAGGCAAAAT	GTGTTGAAGT	GGCCTATGAA	GTAGCAATTT	3300	
	TCATATAGAA	TATAGTGTGA	AATAAATGTG	TGTGTGTATA	TTATTATTA	TCAATGCAAT	3360	
	ATTTAAATG	AAATGAGAAC	AAAGAGGAAA	ATGGTAAAAA	CTTGAATGA	GGCTGGGGTA	3420	
	TAGTTTGTCC	TACATAGAAA	AAAAGAGAGA	GCTTCTTAGG	CCTGGGCTCT	TAAATGCTGC	3480	
	ATTATAACTG	AGTCTATGAG	GAAATAGTTC	CTGTCCAATT	TGTTGAATTT	TTTTAAAAAT	3540	
35	GTAATAAAT	TAAACTTTTC	TGGTTTCTGT	GGGAAGGAAA	TAGGGAATCC	AATGGAACAG	3600	
	TAGCTTTGCT	TTGCACTCTG	TTTCAAGATT	TCTGCATCCA	CAAGTTAGTA	GCAAACTGGG	3660	
	GAATACTCGC	TGCAGCTGGG	GTTCCCTGCT	TTTTGGTAGC	AAGGGTCCAG	AGATGAGGTG	3720	
	TTTTTTTGGG	GGAGCTAATA	ACAAAAACAT	TTTAAACTT	ACCTTTACTG	AAGTTAAATC	3780	
	CTCTATTGCT	GTCTTATTC	TCCTTATAG	TGACCAACAT	CTTTTAAAT	TAGATCCAAA	3840	
40	TAAACCATGT	CTCCTAGAGT	TTAGAGGCTA	GAGGGAGCTG	AGGGGAGGAT	CTTACTGAAA	3900	
	GCAOCTGGG	GAGATTGATT	GTCTTAAAC	CTAAGCCCCA	CAAACTTGAC	ACCTGATGAC	3960	
	GTCTGGGAGC	TACAAAATTT	CATTTTCTC	CTCACTGCC	TTCTTCTGAG	TGGCATGGC	4020	
	CTGAATCAAG	GAAAGCCAGG	CCTTGTGGGC	CCCCCTCTT	CGGCTTCTG	CTAAAGCAAC	4080	
	ACCTCCAGCA	GAGATTCCCT	TAAAGTACTC	CAGGTTTTCC	ACCATCTCTC	AGCGTGAATT	4140	
45	AATTTTTAAT	CAGTTTGCTT	TCTCCAGAGA	AATTTTAAAA	TAAATGAAGA	AATAGAAATT	4200	
	TTGAAATGTAT	AAAAGAAAAA	GATCAAGTTG	TCATTTTGA	ACAGAGGGAA	TTTGGGAGA	4260	
	AAGCAGCCCA	AGTAGGTTAT	TTGTACAGTC	AGAGGGCAAC	AGGAAGATGC	AGGCCCTCAA	4320	
	GGGCAGGAG	AGGCCACAAG	GAATATGGGT	GGGAGTAAAA	GCAACATGCT	CTGCTTCATA	4380	
	CTTTTCTCTA	GGCTTGGCAC	TGCCCTTTCC	TTTCTCAGGC	CAATGGCAAC	TGCCATTGTA	4440	
50	GTCCGGTGAG	GGATCAGCCA	ACCTCTTCTC	TATGGCTCAC	CTTATTGGA	GTGAGAAATC	4500	
	AAGGAGACAG	AGCTGACTGC	ATGATGAGTC	TGAAGGCATT	TGCAGGATGA	GCCTGAACCTG	4560	
	GTTGTGCGAG	ACAAACAAGG	CATTTCATGGG	AATTGTTGTA	TTCTTCTGCG	AGCCCTCCTT	4620	
	CTGGGCACTA	AGAAGGCTTA	TGAATTAAT	GCCTATCTAA	AATTCTGATT	TATCTCTACA	4680	
	TTTTCTGTTT	TCCTAATTGA	CCCTAAAAATC	TATGTGTTT	AGACTTAGAC	TTTTTATTGC	4740	
55	CCCCCCCCC	TTTTTTTTTG	AGACGGAGTC	TCGCTCTGAC	GCACAGGCTG	GAGTGCAGTG	4800	
	GCTCGATCT	CTGCTCACTG	AAAGCTCCGC	CTCCGGGTT	CATGCCATTC	TCCTGCCCTA	4860	
	GCCTCCTGAG	TAGCTGGGAC	TACAGGCGCC	CACCACCACG	CCCGGCTAAT	TTTTTGATT	4920	
	TTTAATAGAG	ACGGGGTTTC	ACTGTGTTAG	CCAGGATGGT	CTCGATCTCC	TGACCTCGTG	4980	
	ATCCGCTGCT	CTCGGCTTCC	CAAAGTGCTG	GGATTACAGG	CATGACCCAC	CGCTCCCGGC	5040	
60	CTTGTTTTCC	GTTTAAAGTC	GTCTTCTTTT	AATGTAATCA	TTTTGAACAT	GTGTGAAAGT	5100	
	TGATCATACG	AATGGAATCA	ATCTTGAAAT	ACTCAACCAA	AAGACAGTCG	AGAAGCCAGG	5160	
	GGGAGAAAGA	ACTCAGGGCA	CAAAATATTG	GTCTGAGAA	GGAATTTCTC	GTAAAGCCTAG	5220	
	TTGCTGAAAT	TTCTGCTGTG	AACCAGAAGC	CAGTTTATC	TAAOCCGTAC	TGAAACACCC	5280	
	ACTGTGTTTT	GCTCACTCCC	TCACCTCACG	ATCAAAACCT	GCTACCTCCC	CAAGACTTTA	5340	
65	CTAGTGCCGA	TAAACTTTCT	CAAGAGACAA	CCAGTATCAC	TTCCCTGTTT	ATAAAACCTC	5400	
	TAAACATCTC	TTTGTCTTTT	GAACATGCTG	AAAACCACT	GGTCTGCATG	TATGCCCGAA	5460	
	TTTGTAAATC	TTTTCTCTCA	AATGAAAAT	TAAATTTAGG	GATTCAATTC	TATATTTTCA	5520	
	CATATGTAGT	ATTATTATTT	CCTTATATGT	GTAAGGTGAA	ATTATGTTA	TTTGAATGTTG	5580	
	CAAGAAAATA	TATTTTAAAT	GCTTTCATTT	TTCCCCAGT	GAATGATTTA	GAATTTTTTA	5640	
70	TGTAATATA	CAGAAATGTT	TTTCTTACTT	TTATAAGGAA	GCAGCTGTCT	AAAATGCAAT	5700	
	GGGGTTTGT	TTGCAATGTT	TTAAACAGAG	TTTTAGTATT	GCTATTAATA	GAAGTTACTT	5760	
	TGCTTTTAAA	GAAACTTGGC	TGCTTAAAT	AAGCAAAAAT	TGGATGCATA	AAGTAATATT	5820	
	TACAGATGTG	GGGAGATGTA	ATAAAACAAT	ATTAACTTGG	TTTCTTGT	TGCTGTATT	5880	
	TAGAGATTAA	ATAATTCTAA	GATGATCACT	TTGCAAAAT	ATGCTTATGG	CTGGCATGGA	5940	
75	AATAGAAATA	CTCAATATG	TCTTGTGTGT	ATTAAATGGG	AATATTTTGG	ACAATGTTTC	6000	
	ATTATCAAT	TGTGACATC	ATTAAATAT	ATTGTAATGT	TGGGAAGAGA	TCACTATTTT	6060	
	GAAGCACAGC	TTTACAGATG	AGTATCTATG	ATACATATGT	ATAATAAAT	TGATCGGGT	6120	
	ATTAAAGTA	TTAGAAGGTG	GTTATAATTG	CAGAGTATTC	CATGAATAGT	ACACTGACAC	6180	
	AGGGGTTTTA	CTTTGAGGAC	CAGTGTAGTC	AAGGGAAAAC	ATGAGTTAAA	AAGAAAAGCA	6240	
80	GGCAATATTG	CAGTCTTGAT	TCCTGCCACT	ACAGGATAGA	TAAAGCTGTA	ACTTTAATGA	6300	
	CAAGATGATC	GACCAATAAA	GGTGCTCTGT	GCTTCACAGT	GAATCTTTTC	CCCATGCAGG	6360	
	AGTGTGCTCC	CCTACAAACG	TTAAGACTGA	TCATTTCAAA	AATCTATTAG	CTATATCAAA	6420	
	AGCCTTACAT	TTTAATATAG	GTTGAACCAA	AATTTCAATT	CCAGTAACCT	CTATTGTAA	6480	
	CATTATTTTT	GTGTATGTCT	TCAAGAAATG	TCAATGGATT	TTTGTGTTGA	ATAGTAAAA	6540	
85	ACCGGATACA	TTTCACTGTG	CCTTCAGTAT	TGATTTGGTT	GAATATTGGG	TCATAATGGT	6600	
	TGAGAAGCAT	GGACACTAGA	GCCAGAATGC	TTGGATATGA	ATCCCTGATC	TGTCACCTAC	6660	
	TTCTGTGTGA	CCTTTGAAAG	GCTACTTATT	TCCTCTCTTA	GCTTTCTCAT	TAAATCAAT	6720	
	GAACAAATGCC	AGCCTCATGG	GGTGTGTGAA	TGATTAATTT	AGTTAATATA	CCTAAAGTAC	6780	

ATAGAACACT GCCTGCACAT AGTAAAAGAA TTATAAGTGT GAGGTAGTTG GTAAAAATTAT 6840
GTAGTTGGAT ATACTACCGA ACAATATCTA ATCTCTTTTT AGGGAAATAA AGTTTGTGCA 6900
TATATATAAT CCGGAACAT G

Seq ID NO: 32 Protein sequence:
Protein Accession #: NP_001932.1

1 11 21 31 41 51
MAAAGPRRS RVAVCLHLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECFRS 60
ADLIRSSDDP FRVLNDGSVY TARAVALSDK KRSFTIWLSD KKKQTQKEVT VLEHQKKVS 120
KTRHTRTVL RRAKRRWAPI PCSMQENSLG PFPLFLQOVE SDAAQNYTVF YSISGRGVDK 180
EPLNLFPIYR DTGNLPCTRP VDREYDVFD LIAYASTADG YSADLPLPLP IRVEDENDNH 240
PVFTEAIYNF EVLESSRPGT TVGVVCAIDR DEPDTHMTRL KYSILQOTPR SPGLFSVHPS 300
TGVIITVSHY LDREVVDKYS LIMKVQMDMG QFFGLIGTST CIITVDSND NAPTFRQAY 360
EAFVEENAFN VEILRIPIED KDLINTANWR VNFITLKONE NGHFKISTDK ETNEGVLVSV 420
KPLNYEENRQ VNLEIGVWNE APFARDIPRV TALNRALVTV HVRDLDEGPB CTPAAQYVRI 480
KENLAVGSKI NGYKAYDPEN RENGNGLYKK LHDPKGWITI DEISSGIIS KIILDRSVETP 540
KNELYNITVL AIDKDDRSCT GTLAVNIEDV NDNPEILQE YVVICPKMG YTDILAVDPD 600
EPVHGAPFFY SLPNTSPEIS RLWSLTQVND TAARLSYQKN AGFQETIPI TVKDRAGQAA 660
TKLLRVNLCE CTHTPTCRAT SRSTGVILGK WAILAILLGI ALLFSVLLTL VCGVFGATKG 720
KRFPEDLAQQ NLIISNTEAP GDDRVCSSANG FMTQTNNSS QGFCGTMSGG MKNGGQETIE 780
MKGGNQTLLE SCRGAGHHHT LDSCRGGHTE VDNCRYTYSE WHSFTQPRLG EKLHRCNQNE 840
DRMPSQDVL TYNYEGRGSP AGSVGCCSEK QEEDGLDFLN NLEPKFITLA EACTKR

Seq ID NO: 33 DNA sequence
Nucleic Acid Accession #: Bos sequence
Coding sequence: 64-2583

1 11 21 31 41 51
GGCAGGTCTC GCTCTCGGCA CCTCCCGGCG GCCCGCGTTC TCCTGGCCCT GCCCGGCATC 60
COGATGGCGG CCGCTGGGCG CCGCGGCTCC GTGCGCGGAG CCGTCTGCCT GCATCTGCTG 120
CTGACCCCTCG TGATCTTCAG TCGTGATGGT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180
CCTTCTAATC TAGAGGSCAG CAAATAAAT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240
TCTGCAGACC TCATCCGCTC AAGTGATCCT GATTTCAGAG TTCTAATGA TGGGTGAGT 300
TACACAGCCA GGGCTGTGTC CTTGTCTGAT AAGAAAAGAT CATTTACCAT ATGGCTTTCT 360
GACAAAAGGA AACACACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420
TCGAAGACAA GACACACTAG AGAAACTGTT CTCAGGCGTG CCAAGAGGAG ATGGGCACCT 480
ATTCTCTGCT CTATGCAAGA GAATTCCTTG GGCCTTTTCC CATTTGTTCT TCAACAAGTT 540
GAATCTGATG CAGCAGACAA GTTACTGTC TTCTACTCAA TAAGTGGACG TGGAGTTGAT 600
AAAGAACCTT TAAATTTGTT TTATATAGAA AGACACACTG GAAATCTATT TTGCACTCGG 660
CCCTGCGGAT GTGAAGAATA TGATGTTTTT GATTGTGATG CTATGCGCTC AACTGCGAGT 720
GGATATTGAG CAGATCTGCC CCTCCCACTA CCGATCAGGG TAGAGGATGA AATGACAAAC 780
CACCCTGTTT TCACAGAGAG AATTTATAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840
ACTACAGTGG GGGTGTGTTG TGCCACAGAG AGAGATGAAC CGGACACAAAT GCATACGGCG 900
CTGAAATACA GCATTTTGCA GCAGACACCA AGGTCACTCG GGCCTTTTTC TGTGCACTCC 960
AGCAGAGGCG TAATCACCAC AGTCTCTCAT TATTGGGACA GAGAGGTTGT AGACAAGTAC 1020
TCATTGATAA TGAAGATACA AGACATGGAT GGCAGTTTTT TTGGATTGAT AGGCACATCA 1080
ACTTGATACA TAACAGTAAC AGATTCAAAT GATAATGCAC CCACTTTCAG ACAAATGCT 1140
TATGAAGCAT TTGTAGAGGA AATGCAATC AATGTGAAA TCTTACGAAT ACCTATAGAA 1200
GATAAGGATT TAATTAACAC TGCCAATTGG AGAGTCAATT TTACCATTTT AAAGGGAAT 1260
GAAAATGGAC ATTTCAAATG CAGCAGACAG AAAGAAACTA ATGAAGGTGT TCTTCTGTT 1320
GTAAAGCCAC TGAATATGTA AGAAAACCGT CAAGTGAACC TGGAAATGG AGTAAACAAAT 1380
GAAGCGCCAT TTGTAGAGA TATTCCAGAG GTGACAGCCT TGAACAGAGC CTGTGTTACA 1440
GTTCTATGTA GGGATCTGGA TGAGGGGCGT GAATGCACTC CTGAGGCCCA ATATGTGCGG 1500
ATTAAAGAAA ACTTAGCAGT GGGGTCAAAG ATCAACGCTC ATAAGGCATA TGACCCCGAA 1560
AATAGAAATG GCATGTGTTT AAGGTACAAA AAATGCAATG ATCCTAAAGG TTGGATCACC 1620
ATTGATGAAA TTTCAGGGTC AATCATAACT TCCAAAATCC TGGATAGGGA GGTGAAATC 1680
CCCAAAATAG AGTTGTATAA TATTACAGTC CTGGCAATAG ACAAAGATGA TAGATCATGT 1740
ACTGGAACAC TTGCTGTGAA CATTGAAGAT GTAAATGATA ATCCACCAGA AATACTTCAA 1800
GAATATGTAG TCATTTGCAA ACCAAAATG GGGTATACCG ACATTTTAGC TGTGTATCCT 1860
GATGAACCTG TCCATGGAGC TCCATTTTAT TTCAGTTTGC CCAATACTTC TCCAGAAATC 1920
AGTAGACTGT GGAGCCTCAG CAAAGTTAAT GATACAGCTG CCGTCTTTC ATATCAGAAA 1980
AATGCTGGAT TTCAAGATA TACCATTCTC ATTACTGTAA AAGACAGGGC CGGCCAAGCT 2040
GCAACAAATC TATTAGAGT TAATCTGTGT GAATGTACTC ATCCAACTCA GTGTGCTGG 2100
ACTTCAAGGA GTACAGGAGT AATACTTGA AATGGGCAA TCCTTGCAAT ATTACTGGGT 2160
ATAGCACTGC TCTTTTCTGT ATTGCTAACT TTAGTATGTG GAGTTTTTGG TGCAACTAAA 2220
GGGAAACGTT TTCTGAAGA TTTAGCACAG CAAAACCTAA TTATATCAAA CACAGAAGCA 2280
CCTGAGACG ATAGAGTGTG CTCTGCCAAT GGATTTATGA CCAAACTAC CAACAACCTC 2340
AGCCAAGGTT TTTGTGGTAC TATGGGATCA GGAATGAAA ATGGAGGGCA GGAAACCAAT 2400
GAAATGATGA AAGGAGGAAA CCAGACCTTG GAATCCTGCC GGGGGGCTGG GCATCATCAT 2460
ACCCCTGGAT CCTGCAGGGG AGGACACAGC GAGGTGGACA ACTGCAGATA CACTTACTCG 2520
GAGTGGCACA GTTTTACTCA ACCCGTCTC GGTGAAGAAT CCAATAGAGG ACACACTGGT 2580
TAAAAATTA ACATAAAAGA AATTGCATCG ATGTAATCAG AATGAAGACC GCATGCCATC 2640
CCAAGATTAT GTCTCACTT ATAACTATGA GGAAGAGGA TCTCCAGCTG GTTCTGTGGG 2700
CTGCTGCAGT GAAAAGCAGG AAGAAGATGG CCTTGACTTT TTAATAATTT TGAACCCAA 2760
ATTTATTACA TTAGCAGAG CATGCACAAA GAGATAATGT CACAGTGCTA CAATTAGGTC 2820
TTTGTACGAC GTTTTGGAGG TTTCCAAAAA TAATATTGTA AAGTTCAAAT TCAACATGTA 2880
TGATATGATG GATTTTTTTC TCAATTTTGA ATTATGCTAC TCACCAATTT ATATTTTAA 2940
AGCCAGTTGT TGCTTATCTT TTCCAAAAAG TGAATAATGT TAAACACAG AACTGGTAAA 3000
TCTCAAACTC CAGCACTGGA ATTAAGGTCT CTAAGGCATC TGCTCTTTTT TTTTTTACG 3060
GATATTTTGA TAATAAATAT GCTGGATAAA TATTAGTCCA ACAATAGCTA AGTTATGCTA 3120
ATATCACTT ATTAGTATT CACTTTAAGT GATAGTTTAA AAAATAAACA AGAAATATTG 3180
AGTATCACTA TGTGAAGAAA GTTTTGGAAA AGAAACAATG AAGACTGAAT TAAATTAATA 3240
ATGTTGCAGC TCATAAGAAA TTGGGACTCA CCCCTACTGC ACTACCAAT TCATTTGACT 3300

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85

TTGGAGGCAA AATGTGTGA AGTGCCCTAT GAAGTAGCAA TTTCTATAG GAATATAGTT 3360
 GGAATAAAT GTGTGTGTGT ATATTATTAT TAATCAATGC AATATTTAAA ATGAAATGAG 3420
 AACAAAGAGG AAAATGGTAA AAACCTGAAA TGAGGCTGGG GTATAGTTTG TCCTACAATA 3480
 GAAAAAGAG AGAGCTTCCT AGGCCTGGGC TCTTAAATGC TGCATTATAA CTGAGTCTAT 3540
 GAGGAAATAG TTCTGTGCA ATTTGTGTAA TTTGTTTAAA ATTGTAATAA AATTAAACTT 3600
 TTCTGTGTTT TGTGGGAAGG AAATAGGGAA TCCAATGGAA CAGTAGCTTT GCTTGCAGT 3660
 CTGTTTCAAG ATTTCTGCAT CCACAAGTTA GTAGCAAACT GGGGAATACT CGCTGCAGCT 3720
 GGGGTTCCCT GCTTTTGTGT AGCAAGGGTC CAGAGATGAG GTGTTTTTTT CGGGGAGCTA 3780
 ATAACAAAAA CATTTTAAAA CTACCTTTTA CTGAAGTTAA ATCTCTTATT GCTGTTTCTA 3840
 TTCTCTCTTA TAGTGACCAA CATCTTTTAA ATTTAGATCC AAATAACCAT GTCCCTCTAG 3900
 AGTTTAGAGG CTAGAGGGAG CTGAGGGGAG GATCTTACTG AAAGCACCCCT GGGGAGATTG 3960
 ATTGTCTCTA AACCTAAGCC CCAAAACTT GACACCTGAT CAGGTCTGGG AGCTACAAAA 4020
 TTTTATTTTT CTCTCACTG CCCTTCTTCT GAGTGGCATT GGCCTGAATC AAGGAAAGCC 4080
 AGGCCTTGTG GGGCCCTTTC TTTGCGCTTT CTGCTAAAGC AACACCTCCA GCAGAGATTG 4140
 CCTTAAGTGA TCTCAGGTTT TCCACCATCC TTCAGGTTGA ATTAATTTTT AATCAGTTTG 4200
 CTTTCTCCAG AGAAATTTTA AAATAATAGA AGAAATAGAA ATTTTGAATG TATAAAGAA 4260
 AAAGATCAAG TTGTCAATTT AGAACAGAGG GAACCTTGGG AGAAAGCAGC CCAAGTAGGT 4320
 TATTGTGACA GTCAAGGGG ACAGGGAAGA TGCAAGGCTT CAAGGGCAAG GAGAGGCCAC 4380
 AAGGAATATG GGTGGGAGTA AAAGCAACAT CGTCTGCTTC ATACTTTTTC CTAGGCTTGG 4440
 CACTGCCCTT TCTTCTCTCA GGCCAATGGC AACTGCCATT TGAGTCCGTT GAGGAGTCAG 4500
 CCAACCTCTT CTCTATGGCT CACCTTATTT GGAGTGAGAA ATCAAGGAGA CAGAGCTGAC 4560
 TGCATGATGA GTCTGAAGGC ATTTGCAGGA TGAGCCTGAA CTGTTGTGTC AGAACAAACA 4620
 AGGCATTCTT GGAATTTGTT GTATTCTTCT TGCAGCCCTC CTTCCTGGGC CTAAGAAGGT 4680
 CTATGAATTA AATGCCATCT TAAATTTCTG ATTTATTTCT ACATTTTCTG TTTTCTAATT 4740
 TGACCCCTAA ATCTATGTGT TTTAGACTTA GACTTTTAT TTGCCCCCCC CCCTTTTTTT 4800
 TTGAGACGGA GTCTGCTCT GACGCACAGS CTGGAGTGCA GTGGCTCCGA TCTCTGCTCA 4860
 CTGAAAGCTC CGCTCCCGG GTTCATGCCA TTCTCTGCTC TCAGCCTCCT GAGTAGCTGG 4920
 GACTACAGGC GCCCACCACC ACGCCCGGCT AATTTTGTGT ATTTTAAATA GAGACGGGGT 4980
 TTCACTGTGT TAGCCAGGAT GGTCTCGATC TCCTGACCTC GTGATCCGCC TGCCCTCGGC 5040
 TCCCAAAGTG CTGGGATTAC AGGCATGACC CACGCTCCCG GGCCTGTTT TCCGTTTAAA 5100
 GTCGTCTTCT TTTAATGTAA TCATTTTGAA CATGTGTGAA AGTTGATCAT ACGAATTTGA 5160
 TCAATCTTGA AATACTCAAC CAAAGACAG TCGAGAAGCC AGGGGAGAG AGAAGTCAAG 5220
 GCACAAATA TTGCTCTGAG AATGGAATTC TCTGTAAGCC TAGTTGCTGA AATTCTCTGC 5280
 TGTAACAGCA AGCCAGTTTT ATCTAACGGC TACTGAAACA CCCACTGTGT TTTGCTCACT 5340
 CCCACTACC GATCAAAACC TGCTACCTCC CCAAGACTTT ACTAGTCCCG ATAAACTTTC 5400
 TCAAGAGCA ACCAGTATCA CTTCCTGTT TATAAAACCT CTAACCATCT CTTTGTCTCT 5460
 TGAACATGCT GAAAGCACC TGGTCTGCT GTATGCCGA ATTTGTAATT CTTTCTCTC 5520
 AAATGAAAT TTAATTTTAG GGATTCATT CTATATTTT ACATATGTAG TATTATTATT 5580
 TCCTTATATG TGTAAGGTGA AATTATGTT ATTTGAGTGT GCAAGAAAT ATATTTTAA 5640
 AGCTTTCATT TTTCCCGAG TGAATGATT AGAATTTTT ATGTAAATAT ACAGAAATGT 5700
 TTTTCTTACT TTTATAAGA AGCAGCTGTC TAAATGTCAG TGGGGTTTGT TTTGCAATGT 5760
 TTTAAACAGA GTTTTAGTAT TGCTATTAA AGAAGTTACT TTGCTTTTAA AGAACTTGG 5820
 CTGCTTAAAA TAAGCAAAAA TTGGATGCAT AAAGTAATAT TTACAGATGT GGGGAGATGT 5880
 AATAAACA TAATAACTTG GCTGCTTAAA ATAGCAAAA ATTGGATGCA TAAAGTAATA 5940
 TTTACAGATG TGGGGAGATG TAATAAACA ATATTAACT GGTTTCTTGT TTTTGTGTA 6000
 TTTAGAGATT AAATAATCT AAGATGATCA CTTTGCAAAA TTATGCTTAT GGCTGGCATG 6060
 GAAATAGAAA TACTCAATTA TGCTTTTGT GTATTAATG GGAATATTT GGACAATGTT 6120
 TCATTATCAA ATGTGACACA TCATTAAAT ATATTGTAAT GTTGGGAAGA GATCACTATT 6180
 TTGAAGCACA GCTTTACAGA TGAGTATCTA TGATACATAT GTATAATAA TTTTGTATCG 6240
 GTATTAAAG TATTGAAGG TGGTTATAAT TGCAGAGTAT TCCATGAATA GTACACTGAC 6300
 ACAGGGGTT TACTTTGAGG ACCAGTGTAG TCAAGGGAAC ACATGAGTTA AAAGAAAAAG 6360
 CAGGCAATAT TGCAGCTTG ATTCTGCCAC TTACAGGATA GATAATGCTT GAACCTTAA 6420
 GACAAAGTGA TCCAAACATA AAGGTGCTCT GTGCTTCACA GTGAATCTTT TCCCATGCA 6480
 GGAGTGTGCT CCCCTACAAA CGTTAAGACT GATCATTTC AAAATCTATT AGCTATATCA 6540
 AAAGCTTATC ATTTTAAAT AGGTTGAACC AAAATTTCAA TTCCAGTAAC TTCTATTGTA 6600
 ACCATTATTT TTGTGTATGT CTCAAGAAT GTTCATTGGA TTTTGTGTT TAATAGTAAA 6660
 ATACCGGATA CAITTCACGT GTCCTTCAGT ATTGATTGG TTGAATATTG GGTCATAATG 6720
 GTTGAGAAGC ATGGACACTA GAGCCAGAA GCTTGGATAT GAATCTGGA TCTGTCACTT 6780
 ACTTCTGTGT GACCTTTGAA AGGCTACTTA TTTCTCTCT TAGCTTCTC ATTAAAAATCA 6840
 ATGAACAATG CCGACTCAT GGGGTTGTTG AATGATTAAA TTAGTTAATA TACCTAAAGT 6900
 ACATAGAACA CTGCTGCAC ATAGTAAAG AATTATAAGT GTGAGGTAGT TGGTAAATTT 6960
 ATGTAGTTGG ATATACTACC GAACAATATC TAATCTCTTT TTAGGGAAT AAAGTTTGTG 7020
 CATATATATA ATCCGAAAC ATG

Seq ID NO: 34 Protein sequence:
 Protein Accession #: NP_077741.1

1 11 21 31 41 51
 70
 75
 80
 85

MAAGPRRSV RGAVCLHLL TLVIFSRDGE ACKKVLNV SKLEADKIIG RVNLEECFRS 60
 ADLIRSSDPD FRVNDGVS TARAVLSDK KRSFTIWLSD KRKQTQKEVT VLEHOKKVS 120
 KTRHRTETVL RRAKRRWAPI PCSMQENSIG PFPFLQQVE SDAQNYTVF YSISGRGVND 180
 EPLNLFYIER DTGNLFCTRP VDREYDVFD LIAYASTADG YSADLPLPLP IRVEDENDNH 240
 PVFTEAIYF EVLESSRPGT TVGVVCAADR DEPDIMHRL KYSILQOTPR SPGLFSVHPS 300
 TGVITTVSHY LDREVDKYS LIMKVQDMQG QFFGLIGTST CIITVDSND NAPTFRQAY 360
 EAFVEENAFN VEILRIPIED KDLINTANWR VNFILKNGNE NGHFKISTDK ETNEGVLVV 420
 KPLNYEENRQ VNLGIVNNE APFARDIPRV TALNRALVTV HVRLDEGPE CTPAAQYVRI 480
 KENLAVGSKI NGYKAYDPEN RGNGLRYKK LHPKGWITI DEISGSIITS KILREVETP 540
 KNELYNITVL AIDRDRSCT GTLAVNIEDV NDNPEILQE YVVICPKPM YTDILAVDP 600
 BPHVGAFFYF SLNPTSPTS RLWSLTKVND TAARLSYQKN AGFOEYTIPI TVKDRAGQAA 660
 TKLLRVNLCE CTHPTQCRAT SRSTGVILGK WAILAILLGI ALLEFVLLTL VCGVFGATKG 720
 KRFPEDLAQQ NLINISNTEP GDDRVCISANG FMTQTNNSS QGFCGTMGSG MKNGGQETIE 780
 MMKGNGQTLB SCRGAGHHT LDSCRGGHTS VDNCRITYSE WHSFTQPRLG EESIRGHTG

Seq ID NO: 35 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 146-1273

1 11 21 31 41 51
GGGAGTGGGC GTGGCGGTGC TGCCCAAGTG AGCCACCGCT GCTTCTGCCC AGACACGGTC 60
5 GCCTCCACAT CCAGGTCCTT GTGCTCCTCG CTTCCTGTT CCTTTCCAC GCATTTTCCA 120
GGATAACTGT GACTCCAGGC CCGCAATGGA TGCCCTGCAA CTAGCAAAAT CGGCTTTTGC 180
CGTTGATCTG TTCAACAAC TATGTGAAAA GGAGCCACTG GGCATATGTC TCTTCTCTCC 240
AATCTGTCTC TCCACCTCTC TGTCACCTGC TCAAGTGGGT GCTAAAGGTG ACATCTGCAA 300
TGAAATGGA CAGGTCTTTC ATTTTGAAAA TGTCAAAGAT ATACCCCTTG GATTTCAAAC 360
10 AGTAACATCG GATGTAAACA AACTTAGTTC CTTTACTCA CTGAAACTAA TCAAGCGGCT 420
CTACGTAGAC AAATCTCTGA ATCTTCTAC AGAGTTCTAT AGCTCTACGA AGAGACCCCTA 480
TGCAAGGAA TTGGAACCTG TTGACTTCAA AGATAAATG GAAGAAACGA AAGGTCTAGT 540
CAACAACCTA ATTAAGGATC TCACAGATGG CCACTTTGAG AACATTTTAG CTGACAACAG 600
TGTGAACGAC CAGACCAAAA TCCTTGTGGT TAATGCTGCC TACTTTGTTG GCAAGTGGAT 660
15 GAAGAAATTT CTTGAATCAG AAAAAGATGAT TGATCCCAAG GCTTGTCTG AAAATCTAGG 720
CAAAACAGTG CAGATGATGA ACATGGAGGC CACGTTCTGT ATGGGAAACA TTGACAGTAT 780
CAATTGTAAG ATCATAGAGC TTCTTTTCAA AAATAAGCAT CTCAGCATGT TCATCTCTACT 840
ACCAAGGAT GTGAGGATG AGTCCACAGG CTTGGAGAAG ATTGAAAAC AACTCAACTC 900
AGAGTCACTG TCACAGTGGG CTAATCCAG CACCATGGCC AATGCCAAG TCAAACCTCTC 960
20 CATTCACAAA GTTGAAGTGC AAAAGATGAT TGATCCCAAG GCTTGTCTG AAAATCTAGG 1020
GCTGAACAT ATCTTCAGTG AAGACACATC TGATTTCTCT GGAATGTCTG AGACCAAGGG 1080
AGTGGCCCTA TCAAGTGGT TCCACAAAGT GTGCTTAGAA ATAAGTGAAG ATGGTGGGGA 1140
TTCCATAGAG GTGCCAGGAG CACGATCCT GCAGCACAAG GATGAATGA ATGCTGACCA 1200
TCCTTTTATT TACATCATCA GGCACAACA AACTCGAAAC ATCATTTTCT TTGGCAAAAT 1260
25 CTGTTCTCTC TAAGTGGCAT AGCCCATGTT AAGTCCCTCC TGACTTTTCT GTGGATGCCG 1320
ATTTCTGTAA ACTCTGCATC CAGAGATTC TTTTCTAGAT ACAATAAAT GCTAATGTG 1380
CTGGATCAGG AAGCCGCGAG TACTTGTCTC ATGTAGCCTT CACACAGATA GACCTTTTCT 1440
TTTTTCCAAT TCTATCTTTT GTTTCCTTTT TTCCATAAG ACAATGACAT ACGCTTTTAA 1500
TGAAAAGGAA TCACGTTAGA GGAATAATAT TTATTCATTA TTTGTCAAAT TGTCCGGGTG 1560
30 AGTTGGCAGA AATTAAGTCA TCCACAAAGA AAATTCCTAT AAGGAAGATT TGAAGCTCTC 1620
TCTTCCAGC ACTATGCTTT CCTTCTTTGG GATAGAGAAT GTTCCAGACA TTCTCGCTTC 1680
CCTGAAGGAG TGAAGAAGT GTAGTGCATG GGACCCACGA AACTGCCCTG GCTCCAGTGA 1740
AACTTGGGCA CATGCTCAGG CTACTATAGG TCCAGAAGTC CTTATGTTAA GCCCTGGCAG 1800
GCAGGTGTTT ATTAATAATC TGAATTTTGG GGATTTTCAA AAGATAATAT TTTACATACA 1860
35 CTGTATGTTA TAGAATCTCA TGGATCAGAT CTGGGGCAGC AACCTATAAA TCAACACCTT 1920
AATATGCTGC AACAAAATGT AGAATATTC GACAAAATGG ATACATAAAG ACTAAGTAGC 1980
CCATAAGGAG TCAAAATTTG CTGCCAAATG CGTATGCCAC CACTTACAA AACACCTTCG 2040
TTGCGCAGAG TTTTCAGATT GTGGAATGTT GGATAAGGAA TTATAGACCT CTAGTAGCTG 2100
AAATGCAAGA CCCCAAGAGG AAGTTCAGAT CTTAATATAA ATTCACCTTC ATTTTGTATA 2160
40 GCTGTCCCAT CTGCTCATGT GGTGGCCTC AGACTGGTGG CAGGGGCTTC TAGCTGACTC 2220
GCACAGGGAT TCTCACAATA GCCGATATCA GAATTTGTGT TGAAGGAAC TGTCTCTTCA 2280
TCTAATATGA TAGCGGGAAA AGGAGAGGAA ACTACTGCCT TTAGAAAATA TAAGTAAAGT 2340
GATTAAAGTG CTCACGTTAC CTTGACACAT AGTTTTCAG TCTATGGGTT TAGTTACTTT 2400
AGATGGCAGG CATGTAACCT ATATTAAATG TAATTTGTAA AGTTGGGTGG ATAAGCTATC 2460
45 CCGTTTGGCG GTTCTGGAT TACTTCTCTA TAAAAAATAT ATATTACCA AAAAAATTTG 2520
TGACATTCCT TCTCCCATCT CTTCTCTGAC ATGCATTGTA AATAGGTTCT TCTTGTCTG 2580
AGATTCAATA TTGAATTTCT CCTATGCTAT TGACAATAAA ATATTATTGA ACTACC

Seq ID NO: 36 Protein sequence:
Protein Accession #: NP_002630.1

1 11 21 31 41 51
MDALQLANSA FAVDLFKQLC EKEPLGNVLF SPICLSTSLA LAQVGAKGDT ANEIGQVLHF 60
55 ENVDIPPGF QTVSDVNLK SSFYSKLKIK RLYVDKSLNL STEFISSTRK PYAKELETVD 120
FKDKLEETKG QINNSIKDLT DGHFENILAD NSVNDQTKIL VVNAAYFVKG WMKFPESET 180
KECPFLNKAT DTPFQMMNID EATFCMGNID SINCKIIELE FQNKHLSMFI LLPKDVESDE 240
TGLEKIEKQL NSESLSQWTN PSTMANAKVK LSIPKFKVEK MIDPKACLEN LGLKHIPSED 300
60 TSDPSGMSSET KGVALSNIHV KVCLEITEDG GDSIEVPGAR ILQHKDELNA DHPFIYIIRH 360
NKTRNIIFFG KFCSP

Seq ID NO: 37 DNA sequence
Nucleic Acid Accession #: NM_0168583
Coding sequence: 72-842

1 11 21 31 41 51
GGAGTGGGGG AGAGAGAGGA GACCAGGACA GCTGCTGAGA CCTCTAAGAA GTCCAGATAC 60
70 TAAGAGCAAA GATGTTTCAA ACTGGGGGCC TCATTGTCTT CTACGGGCTG TTAGCCCGA 120
CCATGGCCCA GTTTGGAGGC CTGCCCGTGC CCTGGGACCA GACCCCTGCC TTGAATGTGA 180
ATCCAGCCCT GCCCTTGAGT CCCACAGGTC TTGCAGGAAG CTTGACAAAT GCCCTCAGCA 240
ATGGCCTGCT GTCTGGGGGC CTGTTGGGCA TTCTGSAAAA CCTTCCGCTC CTGGACATCC 300
75 TGAAGCTTGG AGGAGGTACT TCTGGTGGCC TCCTTGGGGG ACTGCTTGGG AAAGTGAGCT 360
CAGTGATPCC TGGCCTGAAC AACATCATTG ACATAAAGGT CACTGACCCC CAGCTGCTGG 420
AACTTGGCCT TGTGAGAGC CCTGATGGCC ACCGTCTCTA TGTACCATC CTCTCTGGCA 480
TAAAGCTCAA AGTGAATACG CCCCTGGTGG GTGCAAGTCT GTTGGGCTG GCTGTGAAGC 540
TGGACATCAC TGCAGAAATC TTAGCTGTGA GAGATAAGCA GGAGAGGATC CACCTGGTCC 600
TTGGTGACTG CACCCATTCC CTTGGAAGCC TGCAAAATTC TCTGCTTGAT GGACTTGGCC 660
80 CCTCCCATCA TCAAGGTCTT CTGGACAGCC TCACAGGAT CTTGAATAAA GTCCGCTGCT 720
AGTTGGTTCG GGGCAACGTT TGCCCTCTGG TCAATGAGGT TCTCAGAGGC TTGGACATCA 780
CCCTGGTCCA TGACATTGTT AACATGCTGA TCCACGGACT ACAGTTTGTG ATCAAGGTCT 840
AAGCCTTCCA GGAAGGGGCT GGCCCTCTGT GAGCTGCTTC CCACTGCTCA CAGATGGCTG 900
GCCCATGTGC TGAAGATGTA CACAGTTGCC TTCTCTCGGA GGAACCTGCC CCTCTCTCCT 960
85 TCCACCCAGG CCGTGTAAAC ATCCCATGTG CCTCACCTAA TAAATGGCT CTCTTCTGCT 1020
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 38 Protein sequence:
Protein Accession #: NP_057667

```
5 1 11 21 31 41 51
MFQTGGLIVP YGLLAQTMAG FGGLPVPLDQ TLPLNVNPAI PLSPPTGLAGS LTNALSNGLL 60
SGGLLIGILEN LPPLLDILKPG GGTSGGLLGG LLGKVTSVIP GLNNIIDIIV TDPQLLELGL 120
VQSPDGHRILY VTPIPLGIRLQ VNTPLVGASL LRLAVKLDIT AEILAVRDKQ ERIHLVLGDC 180
THSPGLSLIS LLDGLGLPLPI QGLLDSLTGI LNKVLPPLVQ GNVCPVLNVEV LRGLDITLVH 240
10 DIVNMLIHGL QFVIKIV
```

Seq ID NO: 39 DNA sequence
Nucleic Acid Accession #: NM_004363.1
Coding sequence: 115-2223

```
15 1 11 21 31 41 51
CTCAGGCGAG AGGGAGGAAG GACAGCAGAC CAGACAGTCA CAGCAGCCTT GACAAAAAGT 60
TCTTGGAACT CAAGCTCTTC TCCACAGAGG AGGACAGAGC AGACAGCAGA GACCATGGAG 120
TCTCCCTCGG CCCCTCCCCA CAGATGCTGC ATCCCTGGGC AGAGGCTCCT GCTCAGAGCC 180
TCACTTCTAA CCTTCTGGAA CCCGCCACC ACTGCCAAGC TCACTATTGA ATCCAGCCG 240
TTCAATGTGC CAGAGGGGAA GGAGGTGCTT CTACTTGTCC ACAATCTGCC CCAGCATCTT 300
TTTGGGTACA GCTGTAGCAA AGGTGAAAGA GTGGATGGCA ACCGTCAAAT TATAGGATAT 360
GTAATAGGAA CTCACAAGC TACCCAGGG CCCGCATACA GTGGTCGAGA GATAATATAC 420
CCCAATGCAT CCTGCTGAT CCAGAACATC ATCCAGATG ACACAGGATT CTACACCTTA 480
CAGTCATATA AGTCAGATCT TGTGAATGAA GAAGCAACTG GCCAGTTCCG GGTATACCCG 540
GAGCTGCCCA AGCCCTCCAT CTCACGCAAC AACTCCAAAC CCGTGGAGGA CAAGGATGCT 600
GTGGCCTTCA CCTGTGAACC TGAGACTCAG GAGCACAACCT ACCTGTGGTG GGTAAACAAT 660
CAGAGCTCC CGTCACTGC CAGGCTGCAG CTGTCCAATG GCAACAGGAC CCTCACTCTA 720
TTCAATGTCA CAAGAAATGA CACAGCAAGC TACAAATGTG AAACCCAGAA CCCAGTGAGT 780
GCCAGGCGCA GTGATTCAGT CATCCTGAAT GTCTCTTATG GCCCGATGCC CCCACCAATT 840
TCCCTCTTAA ACACATCTTA CAGATCAGGG GAAATCTGAA ACCTCTCTCG CCACGCAGCC 900
TCTAACCCCA CTGCACAGTA CTCTTGGTTT GTCAATGGGA CTTTCCAGCA ATCCACCCAA 960
GAGCTCTTTA TCCCAACAT CACTGTGAAT AATAGTGGAT CCTATACGTG CCAAGCCCAT 1020
AACTCAGACA CTGGCCTCAA TAGGACCACTA GTCACGACGA TCACAGTCTA TGCAGAGCCA 1080
CCCAACCCCT TCATCACCAG CAACAACCTC AACCCCGTGG AGGATGAGGA TGCTGTAGCC 1140
TTAACTGTG AACCTGAGAT TCAGAACACA ACCTACCTGT GGTGGGTAAA TAATCAGAGC 1200
CTCCCGGTCA GTCCAGGCTG GCAGCTGTCC AATGACAACA GGACCCCTAC TCTACTCAGT 1260
GTCAAGAAGA ATGATGTAGG ACCCTATGAG TGTGGAATCC AGAACGAAAT AAGTGTGAC 1320
CACAGCGACC CAGTCATCCT GAATGTCTCT TATGGCCGAG ACGACCCAC CATTTCCCCC 1380
TCATACACCT ATTACCGTCC AGGGGTGAAC CTCAGCCTCT CCTGCCATGC AGCCTCTAAC 1440
CCACCTGCAC AGTATTCTTG GCTGATTGAT GGGAAACATCC AGCAACACAC ACAGAGGCTC 1500
TTTATCTCCA ACATCACTGA GAAGAACAGC GGACTCTATA CCTGCCAGGC CAATAACTCA 1560
GCCAGTGGCC ACAGCAGGAC TACAGTCAAG ACAATCACAG TCTCTGCGGA GCTGCCCAAG 1620
CCCTCCATCT CCAGCAACAA CTCCAAACCC GTGGAGGACA AGGATGCTGT GGCCTTCAAC 1680
TGTGAACCTG AGGCTCAGAA CACAACCTAC CTGTGTGTGG TAAATGTGTA GAGCCTCCCA 1740
GTCACTCCCA GGCTGCAGCT GTCCAATGGC AACAGGACCC TCACTCTATT CAATGTGACA 1800
AGAAATGAGC CAGAGCCCTA TGTATGTGGA ATCCAGAACT CAGTGAGTGC AAACCCGAGT 1860
GACCCAGTCA CCTGTGATGT CCTCTATGGG CCGGACACCC CCATCATTTT CCCCCAGAC 1920
TCGCTTTACC TTTGGGGAGC GAACCTCAAC CTCTCTGCCC ACTCGGCTCT TAACCCATCC 1980
CCGCAATATT CTGGCGGTAT CAATGGGATA CCGCAGCAAC ACACACAAGT TCTCTTTATC 2040
GCCAAATCA CGCCAAATAA TACCGGGACC TATGCCCTGT TTGTCTCTAA CTGGCTACT 2100
GGCCGCAATA ATTCCATAGT CAAGAGCATC ACAGTCTCTG CATCTGGAAC TTCTCTCTGT 2160
CTCTCAGCTG GGGCCACTGT CGGCATCATG ATTGGAGTGT TGGTGGGGT TGCTCTGATA 2220
TAGCAGCCCT GGTGTAGTTT CTTCATTTCA GGAAGACTGA CAGTTGTTT GCTTCTCTCT 2280
TAAAGCATTT GCAACAGCTA CAGTCTAAAA TTGCTTCTTT ACCAAGGATA TTTACAGAAA 2340
AGACTCTGAC CAGAGATCGA GACCATCCTA GCCAACATCG TGAAACCCCA TCTCTACTAA 2400
AAATACAAAA ATGAGCTGGG CTGTGTGGCG CGCACCTGTA GTCCAGTTA CTCGGGAGGC 2460
TGAGGCGAGA GAATCGCTTG AACCCGGGAG GTGGAGATTG CAGTGAGCCC AGATCGCACC 2520
ACTGCACCTC AGTCTGGCAA CAGAGCAGA CTCCATCTCA AAAAGAAAAG AAAAGAGAC 2580
TCTGACCTGT ACTCTGAAT ACAAGTTTCT GATACCACTG CACTGTCTGA GAATTTCCAA 2640
AACTTTAATG AACTAATCGA CAGCTTCATG AAACGTGTCA CCAAGATCAA GCAGAGAAAA 2700
TAATTAATTT CATGGGACTA AATGAACATA TGAGGATTGC TGATTCTTTA AATGTCTTGT 2760
TTCCAGATT TCAGGAAACT TTTTCTCTT TAAGCTATCC ACTCTACAG CAATTTGATA 2820
AAATATACTT TTGTGAACAA AATTGAGAC ATTTACATTT TCTCCCTATG TGGTCGCTCC 2880
AGACTTGGGA AACTATTCTT GAATATTTAT ATTGTATGGT AATATAGTTA TTGCACAAGT 2940
TCAATAAAAA TCTGCTCTTT GTATAACAGA AAAA
```

Seq ID NO: 40 Protein sequence:
Protein Accession #: NP_004354.1

```
75 1 11 21 31 41 51
MESPSAPPHR WCIPWQRLLL TASLLTFWNP PTTAKLTIES TPFNVAEGKE VLLLHNLPQ 60
HLFGYSWYKG ERVDGNRQII GYVIGTQAT PGPAYSGREI IYPNASLLIQ NIQNDTFY 120
TLEVIKSDLV NEETATQFRV YPELPKPSIS SNNSKPVEDK DAVAFTEPBE TQDATYLVWV 180
NNQSLPVSPR LQLSNGNRTL TLFNVTRNDT ASYKCBQNP VSARRSDSVI LNVLYGPDAP 240
TISPLNTSYR SGENLNLSCH AASNPAAQYS WFNVTGTFQS TQELFIPNIT VNNSGYSYTC 300
AHNSDTGLNR TTVTITTVIA EPPKPFITS NSNPVEDEDA VALTCEPEIQ NTYLVWVNN 360
QSLPVSPRLQ LSNDNRNLTLL LSVTRNDVGP YECCIQNELS VDHSBPVILN VLYGPDPTI 420
SPSYTYRPG VNLSLSCHAA SNPPAQYSWL IDGNIQHTQ ELFISNITEK NSGLYTCQAN 480
NSAGHSRRT VKTITVSDEL PKPSISSMNS KPVEDKDAVA FTCEPEAONT TYLVWVNGQS 540
LPVSPRLQLS NGNRTLTLFN VTRNDARAYV CGIQNSVSAN RSDPVTLDVL YGPDTPIIIS 600
PDSSYLSGAN LNLSCSHASN PSPQYWRIN GIPQHTQVL FIAKITFNIN GTYACFVSNL 660
85 ATGRNNSIVK SITVSAGSTS PGLSAGATVG IMIGVLVGVA LI
```

Seq ID NO: 41 DNA sequence
Nucleic Acid Accession #: NM_006952.1
Coding sequence: 11-793

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
AATCCGACAA	ATGGCGAAG	ACAACTCAAC	TGTTCTGTGC	TTCCAGGGCC	TGCTGATTTT	60
TGGAATGTG	ATTATTGGT	GTGCGGCAT	TGCCCTGACT	GCGGAGTGCA	TCTTCTTTGT	120
ATCTGACCAA	CACAGCCTCT	ACCCACTGCT	TGAAGCCACC	GACAACGATG	ACATCTATGG	180
GGCTGCCTGG	ATCGGCATAT	TGTGGGCAT	CTGCCTCTTC	TGCCTGTCTG	TCTTAGGCAT	240
TGTAGGCATC	ATGAAGTCCA	GCAGGAAAAT	TCTTCTGGCG	TATTTTATTG	TGATGTTTAT	300
AGTATATGCC	TTTGAAGTGG	CATCTGTAT	CACAGCAGCA	ACACAACGAG	ACTTTTTCAC	360
ACCCAACTCT	TTCTTGAAGC	AGATGCTAGA	GAGGTACCAA	AACAACAGCC	CTCCAACAA	420
TGATGACCAG	TGGAAAAACA	ATGGAGTCAC	CAAAACCTGG	GACAGGCTCA	TGCTCCAGGA	480
CAATTGCTGT	GGGCTAAATG	GTCCATCAGA	CTGGCAAAAA	TACACATCTG	CCTTCCGGAC	540
TGAGAATAAT	GATCGTGACT	ATCCCTGGCC	TGCTCAATGC	TGTGTTATGA	ACAACTCTTA	600
AGAACCTCTC	AACCTGGAGG	CTTGTAAACT	AGGCGTGCTT	GGTTTTTATC	ACAACTCAGG	660
CTGCTATGAA	CTGATCTCTG	GTCCAATGAA	CCGACAGGCC	TGGGGGGTTG	CCTGGTTTGG	720
ATTTGCCATT	CTCTGCTGGA	CTTTTGGGT	TCTCTCTGGT	ACCATGTTCT	ACTGGAGCAG	780
AATTGAATAT	TAAGAA					

Seq ID NO: 42 Protein sequence:
Protein Accession #: NP_008883.1

25
30
35
40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
MAKDNSTVRC	FQGLLIFGNV	IIGCCGIALT	AECIFFVSDQ	HSLYPLLEAT	DNDIYGAAM	60
IGIFVIGCLF	CLSVLGIWGI	MKSSRKILLA	YFILMPIVYA	FEVASCITAA	TQDFFTPNL	120
FLKQMLERYQ	NNSPPNNDQ	WKWNGVTKTW	DRLMLQDNCC	GVNGPSDWQK	YTSAFRTENN	180
DADYPMPRQC	CVNMNLKEPL	NLEACKLGVP	GFYENQGCYB	LISGPMNRHA	WGVAMFGFAI	240
LCNTFWVLGG	TMFYWSRIEY					

Seq ID NO: 43 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 83-2605

40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
GCCGGACAGA	TCTGCGCGTA	TCCTGGAGCC	GGCCCAAGTTG	TGAAGTAGGA	GAGCTTTGGG	60
ACCTCTGTCC	CAAGCAAGAG	AGATGAATGG	AGAGTATAGA	GGCAGAGGAT	TTGGACGAGG	120
AAGATTTCAA	AGCTGGAATA	GGGGAAGAGG	TGGTGGGAAC	TTCTCAGGAA	AATGGAGAGA	180
AAGAGAACAC	AGACCTGATC	TGAGTAAAC	CACAGGAAAA	CGTACTTCTG	AACAACCCCC	240
ACAGTTTTTG	CTTTCAACAA	AGACCCACCA	GTCAATGCAG	TCAACATTGG	ATCGATTTCAT	300
ACCATATAAA	GGCTGGAAGC	TTTATTCTC	TGAAGTTTAC	AGCGATAGCT	CTCCTTTGAT	360
TGAGAAGATT	CAAGCATTGG	AAAAATTTT	CACAAGGCAT	ATTGATTTGT	ATGACAAGGA	420
TGAAATAGAA	AGAAAGGGAA	GTATTTTGGT	AGATTTTAAA	GAACTGACAG	AAGGTGGTGA	480
AGTAACCAAC	TTGATACCA	ATATAGCAAC	TGAAGTAAAG	GATGCACCTG	AGAAAACCTT	540
GGCTTGCAATG	GGTTTGGCAA	TACATCAGGT	GTAACTAAG	GACCTTGAAA	GGCATGCCAGC	600
TGAGTTACAA	GGCCAGGAAG	GATGTCTAA	TGATGGAGAA	ACAATGGTAA	ATGTGCCACA	660
TATTCATGCA	AGGGGTGATC	ACTATGAGCC	TTTGACACAG	CTCAAGAAATG	TCAGAGCAAA	720
TTACTATGGA	AAATACATTG	CTCTAAGAGG	GACAGTGGTT	CGTGTCACTA	ATATAAAGCC	780
CTTTTGCAAC	ATGATGGCTT	TTCTTTGTGC	TGATGTGGA	GAAATTCAGA	GCTTTCTCTT	840
TCCAGATGGA	AAATACAGTC	TTCCACAAA	GTGTCTGTG	CCTGTGTGTC	GAGGCAGGTC	900
ATTTACTGCT	CTCCGAGCT	CTCCTCTCAC	AGTTAGGATG	GACTGGCAGT	CAATCAAAAT	960
CCAGGAATTT	ATGCTGTGAT	ATCAGAGAGA	AGCAGGTGG	ATTCCACGAA	CAATAGAAATG	1020
TGAGCTTGT	CATGATCTTG	TGGATAGCTG	TGTCCCGGGA	GACACAGTGA	CTATTACTGG	1080
AATTTGCAAA	GTCTTCAAG	CGGAAGAAGG	TTCTGSAAT	AAGAATGACA	AGTGTATGTT	1140
CCTTTTGAT	ATTGAAGCAA	ATTCTATTAG	TAATAGCAAA	GGACAGAAAA	CAAAGAGTTC	1200
TGAGGATGGG	TGTAAGCATG	GAATGTTGAT	GGAGTTCTCA	CTTAAAGACC	TTTATGCCAT	1260
CCAAGAGATT	CAAGCTGAAG	AAAACTGTT	TAAACTCATT	GTCAACTCGC	TTTGCCCTGT	1320
CATTTTGGT	CATGAACCTG	TAAAGCAGG	TTTGGCATT	GCACTCTTTG	GAGGAAGCCA	1380
GAAATACGCA	GATGACAAAA	ACAGAATTCC	AATTTGGGGA	GACCCCAACA	TCCTTGTGT	1440
TGGAGATCCA	GGCCTAGGAA	AAAGTCAAT	GCTACAGGCA	GCGTGCAATG	TTGCCCAACG	1500
TGGCTGTAT	GTTTGTGGTA	ACACCACGAC	CACCTCTGGT	CTGACGGTAA	CTCTTTCAAA	1560
AGATAGTTCC	TCTGGAGATT	TTGCTTTGGA	AGCTGGTGCC	CTGGTACTGT	GTGATCAAGG	1620
TATTTGTGGA	ATCGATGAAT	TTGATAAGAT	GGGGAATCAA	CATCAAGCCT	TGTTGGAAGC	1680
CATGGAGCAG	CAAGTATTA	GTCTTGTAA	GGCTGGTGTG	GTGTTAGCC	TTCTGCAAG	1740
AACTTCCATT	ATTGCTGCTG	CAAATCCAGT	TGGAGGACAT	TACAATAAAG	CCAAAACAGT	1800
TTCTGAGAA	TTAAAAATGG	GGAGTGCAT	ACTATCCAGA	TTTGATTG	TCTTTATCCT	1860
GTTAGATACT	CCAAATGAGC	ATCATGATCA	CTTACTCTCT	GAACATGTGA	TTGCAATAAG	1920
AGCTGGAAG	CAGAGAACCA	TTAGCAGTGC	CACAGTAGCT	CGATGAATA	GTCAAGATTC	1980
AAATACTTCC	TACTTTGAAG	TAGTTTCTGA	GAAGCCATTA	TCAGAAAGAC	TAAAGGTGGT	2040
TCCTGAGAA	ACAATAGATC	CCATTCCCA	CCAGCTATTG	AGAAAGTACA	TTGGCTATGC	2100
TGGGAGTAT	GTGTACCCAA	GGCTATCCAC	AGAAGCTGCT	CGAGTTCTCT	AAGATTTTTT	2160
CCTTGAGCTC	CGGAAACAGA	GCCAGAGGTT	AAATAGCTCA	CCAATCACTA	CCAGGCAGCT	2220
GGAATCTTTG	ATTCGTCTGA	CAGAGGCAGC	AGCAAGGTTG	GAATTGAGAG	AGGAAGCAAC	2280
CAAGAAGAC	GCTGAGGATA	TAGTGGAAAT	TATGAAATAT	AGCATGCTAG	GAACTTACTC	2340
TGATGAATTT	GGGAACCTAG	ATTTTGAGCG	ATCCACGAT	GGTCTGGAA	TGAGCAACAG	2400
GTCAACAGCG	AAAAGATTTA	TTTCTGCTCT	CAACAACGTT	GCTGAAAGAA	CTTATAATAA	2460
TATATTTCAT	TTTCATCAAC	TTGGCAGAT	TGCCAAAGAA	CTAAACATTC	AGGTTGCTGA	2520
TTTGAAAAAT	TTTATTTGAT	CACATAATGA	CCAGGGTTAC	CTCTTGAAAA	AAGGCCCAAA	2580
AGTTTCCAG	CTTCAAACTA	TGTAAAAGGA	CTTACCAGAG	TTAGGGCCTC	CTGGGTTTAT	2640
TGCAGATTAA	AGCCATCTCA	GTGAAGATAT	GGTGTGACGC	ACAGACAGAC	AGACACACAC	2700
ACACACACAC	ACACACACAC	ACACACACAC	ACACACAGTC	AAATACTGTT	CTCTGAAAAA	2760
TGATGTCCCA	AAAGTATTAT	AATAGGAAAA	AAGCATTAAA	TATAATAAAC	TAATTTAAGA	2820

5
10
15
20
25
30

AGTGATAAAG TCTCCAGATG CAGTAGCTCA CACTGTAATC ACAGTGACTC AGGAGGCTGA 2880
GGTGAGAGGA TTCTCTGAGG CCAGGGTTTCG AGACCAACCT TGGGCAACAT AGCAAGACCC 2940
CATTTCTTAA AAAAAAATTTA AACTTAGCTG GGTATGGTGG CACATGCGTA 3000
TAGTCTCAGC TACTTGTGAG GCTGAGGCAG GAGGATTCTT TGAGCCGAGG AGTTTGAGGT 3060
TACAGTGAGC CACAATCACA CCAATCACTG CACTCCAGCC TGGGCAATAA AGTAACTCTT 3120
GACTCAAAAA AATAAAAAAA ATTGTAGTGG TAGCCATGTG TTAATTGTTA AATAAATCTT 3180
CCAAAGGGCT AAAAGTAAT TACTTATAAA TTTTATATAG TTGTATTTT GACCTGCCTT 3240
TTATATGTAT GAATATTCCA TAGTTTTCGA TATCAGATGT AGGCATACAG ACAAAATACAT 3300
AAACCAATGA ATATATTACA TATTCTGTGT TCCAATAAAA CTTTATTAT GGACACTAAA 3360
ATTGTAATTT CATAAATTT TCCCATGTCA AGAATACAAA ATACTTGAGT TTTGTTTTTA 3420
GCTATTTAAT AATAGTCTCT ATTTATTCCA CAGGCTGTAG TTTGTAGTCT TGCTTGAAC 3480
AATAGAAACA GACTGATTAA GCAGGAGAAG TTTTGTGAAA GAATTTTGT TGGCTCACG 3540
AATTATTAGA AGGCAGGTGA ACCAGGAGGG TAAAGCTTCA GCAGCAATTT GTAAAAACAT 3600
GCCTTAGAAT TGGACTAAGG AAGAAGCTGC TGACACTCCA CTGCCACACA GGGCACTGGA 3660
AGAAAGTGCT GCTGCTTCC TGCCCACTT TGCAGCAGGA ATAGGTAGAA 3720
GAATGCCCCC ACCGSCACCG GAACAGCAAC AAAAGGATTC TGCAATGAGT GCCTCCCTAA 3780
ATTGCTGAAT TCAAAAAAGA AGTTGCATAC AAAGACATCT GATTGAAAAA GGGTATGTTA 3840
TATGCCCTTT TATAGGCTG CTAGGGAGTT TTCTGTGTT TACTTTCAGG TGGTGGGATC 3900
AATAAGACCA GAATTTCTCA TATGTTGTGA GAGGATTCAT ATGTTACAGG GTTGCCAGCC 3960
AAACTATCAA TCACTATATA ATCCAACAAA CACTTTGTAA CATACAAGAA CTCAGGAAT 4020
GTGAACCAAT GTTGAGAAT CTAATAAAAT ACGCTTCCC GCAACGAGG ATGAATGAA 4080
AATGTAATAA AAAAGAACTG GCAGTGTATA TCAGATGTTT AACTATAGGA CCAGAACTAA 4140
GATGTGGAGA CTATTGCCAT AGACCAATAT GTAAATTTT AAGTGAGGAA GGAAAAATCA 4200
GGAATCAAAA GGGGCCAGGT GCAGTGGCTC ACATCTATAA TCCAGAGCT TTGGGAGTTC 4260
GAGGCAGGAG GATCATTGTA AGCCAGTTT GAGACCAGCC TATGCAACAC ATTGAGACCC 4320
TATCTCTACA AAAAATAGAT TAGCTGGGCA CGGTGGTGCA TGCCTATTGT CCTACCTACT 4380
GTGGAGGCTG AAGTAGGAAA TCACCTGAGC CCGAGAGTTT GAGGTTACAG TGAGCTATGA 4440
TTATACCACT GCATCCAGC CTGGCAAGA GAGCAAGACC TTGTCTCT

Seq ID NO: 44 Protein sequence:
Protein Accession #: CAB55276.2

35
40
45
50

1 11 21 31 41 51
MNGEYRGRGF GRGRFQSWKR GRGGGNFSK WRERHRPDL SKTTGKRTSE QTPQFLLSTK 60
TPQSMQSTLD RPIPYKQWKL YPSEVYSDSS PLIEKIQAPE KPFTRHIDLY DKDEIERKGS 120
ILVDFKELTE GGEVNLNLPD IATELDAPE KTLACMLAI HQVLT KDLE HAELQAQEG 180
LSNDGETMVN VPHIHARVYN YEPLTQLKNV RANYYGKYIA LRGTVVVRVN IKPLCTKMAP 240
LCAACGEIQS FPLPDGKYSL PTKCPVPCR GRSFTALRSS PLVTMDWQS IKIQELMSDD 300
QREAGRIPRT IECLELVDLV DSCVPGDTVT ITGIVKVSNA EESGRNKNDK CMFLLYEAN 360
SISNSKQKKT KSSEDGCKHG MLMEFSLKDL YAIQEIQAEE NLPLKLVNSL CPVIFGHELV 420
KAGLALALFG GSQKYADDKN RPIRGDPHI LVVQDPGLGK SQMLQAACNV APRGVVVCN 480
TTTTSLGTVT LSKDSSSGDF ALEAGALVLG DQICGIDEF DKMGNGHQAL LEAMEQQSIS 540
LAKAGVVCSL PARTSIIAAA NPVGGHYNKA KTVSENLMKG SALLSRFDLV FILLDTPNH 600
HDHLLSEHVI AIRAGKQRTI SSATVARMNS QDSNTSVLEV VSEKPLSERL KVVPEGETIDP 660
IPHQLLRKYI GYARQVYVPR LSTEAAARVLQ DFYLELRKQS QRLNSPITR QLESILIRLT 720
EARARLELRE EATKEDAEDI VEIMKYSMLG TYSDEFGNLD FERSQHSGSM SNRSTAKRFI 780
SALANVAERT YNNIFQFHQL RQIAKELNIQ VADFFENFIGS LNDQGYLLKK GPKVYQLQTM

Seq ID NO: 45 DNA sequence
Nucleic Acid Accession #: NM_005416.1
Coding sequence: 149..658

55
60
65
70
75

1 11 21 31 41 51
ACCAGATCCC AGAGGCTGAA CACCTCGACC TTCTCTGCAC AGCAGATGAT CCCTGAGCAG 60
CTGAAGACCA GAAAAGCCAC TAAGACTTTC TGCTTAATTC AGGAGCTTAG AGGATTCTTC 120
AAAGAGTGTG TCCACGATCC TTTGAAGCAT GAGTTCTTAC CAGCAGAAGC AGACCTTTAC 180
CCACCACTCT CAGCTTCAAC AGCAGCAGGT GAAACAACCC AGCCAGCCTC CACCTCAGGA 240
AATATTGTGT CCCACACCA AGGAGCCATG CCACTCAAAG GTTCCACAC CTGGAACAC 300
AAAGATTCCA GAGCCAGGCT GTACCAAGGT CCCTGAGCCA GGCTGTACCA AGGTCCCTGA 360
GCCAGGCTGT ACCAAGGTCC CTGAGCCAGG TTGTACCAAG GTCCCTGAGC CAGGCTGTAC 420
CAAGGTCCCT GAGCCAGGTT GTACCAAGGT CCCTGAGCCA GGCTACACCA AGGTCCCTGA 480
ACCAGGCAGC ATCAAGGTCC CTGACCAAGG CTTTATCAAG TTTCTGAGC CAGGTGCCAT 540
CAAAGTTCCT GAGCAAGGAT ACACCAAGT TCCTGTGCCA GGCTACACAA AGCTACCAGA 600
GCCATGTCTT TCAACGGTCA CTCAGGCCCC AGCTCAGCAG AAGACCAAGC AGAAGTAATT 660
TGGTGACAG ACAAGCCCTT GAGAAGCCAA CCACCAAGAT CTGGACACCC TCTTCCCATC 720
TGTTCTGTGT TCTTAATGT CTGTAGACCT TGTAAATCAG ACATTGTAC CCAAGCCAT 780
AGTCTCTCTC TTATTGTAT CTTAAAAATA CGTACTATAA AGCTTTTGTT CACACACACT 840
CTGAAGAATC CTGTAAGCCC CTGAATTAAG CAGAAGTCT TCAATGCTTT TCTGCTCTTC 900
GGCTGCTCAG GGTTCATCTG AAGATTCGAA TGAAAGAAA TGCATGTTT CTGCTCTTCC 960
CTCATTAAAT TGCTTTAAT TCCA

Seq ID NO: 46 Protein sequence:
Protein Accession #: NP_005407.1

80

1 11 21 31 41 51
MSSYQOKQTF TPPPQLQQQQ VKQPSQPPRQ EIFVPTTKEP CHSKVPQPN TKIPEPGCTK 60
VPEPGCTKVP EPGCTKVPEP GCTKVPEPGC TKVPEPGCTK VPEPGYTKVP EPGSIKVPDQ 120
GFIKPEPGA IKVPEGYTK VPVEGYTKLP EPCPSTVTFG PAQOKTKQK

85
Seq ID NO: 47 DNA sequence:
Nucleic Acid Accession #: Eos sequence

1 11 21 31 41 51
GCGTCGTGTG CAGGCGTCCC CGGGCTGTGG ATAATTAGAC ACGTCTCTCC CTCATTGCCC 60
AAGGCTCGTT AGAATTTCGCC CTAGAGCTGT ATCATGTATT TTCTTTCAAA TTAACITTCG 120
5 TTGCAATTAA GCTTAGGGAA CCAGCAACAA AAGCAAACTT GGCCCGAGGT CGTTCACCGC 180
GAAATGGAT TAGAGAACT TCTTCCCAGA TTTAAGGGGA AAGATTCTGT CGGCCAGCGC 240
TTTGGGGAAA GTGCCCGGAC CGCAGAGGCG ACGCAGGGG AGCAGGAAGC TGCTCAGCGT 300
AGTCGGCGTT GCGCGCAGCG GTGGCCTTCC TCATCTGGGC GATGTGGGCT CCTAGAAGAG 360
10 TAAGGATAAC ATCCTGAAA TGACTTCTGT ACGGTTTGTG CCAACTGCA CACTCATGAC 420
TTGGAGCTGC CCTGTGGAGT TACAGTTTAC CAAACACATT CATGAACATA ATCTCATTTA 480
CTAAAAACTT TGTGAGAAAT TTCTTTTACT AAAATTTTTT CTTATTACAA A

Seq ID NO: 48 DNA sequence:
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51
TTCCAAATTT TTTTITTTGT AATAAGAAAA AATTITAGTA AAAGAAAAT CTCACAAAGT 60
20 TTTTGTATA TGAGATTATG TTCATGAATG TGTITGGTAA ACTGTAACT CACAGGCGAG 120
CTCCAAGTCA TGAGTGTGCA GTTGGGCTCA AACCGTACAG AAGTCATTTC CAGGATGTTA 180
TCCTTACTCT TCTCGAGAGC CACATCGCCC AGATGAGGAA GGCCACCGCT GCGCCCAAGC 240
CCGACTACCG TGAGCAGCTT CCTGCTCCCC TGTCTGTGCC TCTGCGGTGT GGGCACTTTC 300
25 CCCAAGAGCG TGCGCCGAGG AATCTTTCCC CTTAAATCGG GGAAGAAAT TCTCTAATCC 360
ATTTTCGCGG TGAACGACCT CGGGCCAAGT TTGCTTTTGT TGCTGGTTCC CTAAGCTTAA 420
TTGCAAGCAA AGTTAATTGG AAAGAAAATA CATGATACAG CTCTAGGGCG AATTCTAAGC 480
AGCCTTGGGC AATGAGGGAA GAACGTGTCT AGTTATCCAC AGCCCCGGGA CGCTGCACA 540
CGACGCT

Seq ID NO: 49 DNA sequence
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51
TCTTCTCTCT GCTGCTCGTT TGTCTCTCCT GTGCTCTCT TCTTCTTTC CCTGCGCGCT 60
35 CCTGCGGACC TCTGTGTCTCT CTCTCTGAT GGCGGGGGCG GGGAGAAGCT GACCGGTGAG 120
ACCGTAGACC CGAAACCACT GGGTGTGACA AGCCGGTCCG CGCTTTTTF GGGAGAAGCC 180
GACACATGCA GACCACTTTT CCTGGAACNG CATGACCATG TTATTACTAT GGGCCGCTC 240
40 CCCAACCAAA GTGTTTAAAA CTTTTTAGGG CACCCCAAA ATTTTTTTT TTTTTTTTT 300
TTTATTAAA AACTCTAAT ATTTATATTA AATACAAAGA TACCCAAACC CTTTATGCT 360
CTTCTCTGA TCTGTGCTCT TTTCTTTGA CAGCATCTCC ATTTTTTTC TGCTGCTTCA 420
TCGCTGTAGC CATGGGAATC CGTTTCATTA TTATGTAGC AATATGGAGT GCTGTATTCC 480
TAAAGAAACT GACACAGGAG AATCACTTGA ACTTGGGAGG CAGAGTTTGC AGTGAGCCGA 540
GATTGAACCA GTGCACTCCA GCCTTGGCAG CGGAGCAAGA TTCTGTACA GTTCCTGAAG 600
45 TGCTGTATC TGCTGTGACG CCCATCCTCG GTTCCATTGC GCTGCCAGCG AGGCTGCTGG 660
GACGTGGGGA GAGCTGTCT ATATATCOGG GTGAAGCTCA GCTGTGCAC ACCTTGGATG 720
CGGGTCTCT CCTGGCCCCG GGGACCTAGT ATTTTGTCCA CGAGTGTACA CCAACCAAG 780
GAGACAGCAT CATTTATGAG CCTGCAGCAT CCACCTACT GCTGTATCCA GTTTCATTG 840
ACTG

Seq ID NO: 50 DNA sequence
Nucleic Acid Accession #: L05187
Coding sequence: 1991..2260

1 11 21 31 41 51
CTGCAGGGAG GCAGGTAGAA AAGGCTTTTG GGTTTTCAGG TGGGGGSCAG TCTAGCCTGA 60
55 TCAGAAAGGA GGAAGAGGCC AGGGCAGATG TCTGGGTGGA GTGAAGGAA AAGTGTATCC 120
CAGAAGAAAG ATTAGCCCCT GAAAGTCCCT GAAGTAGGAG AAGGGTAAAG GTGTGGTTGG 180
60 TGAAGGAAAG CAGGTTTCC CAGATTAGCA ACCAGTCAGG GGGAGGAAGG TGAGAGTGGG 240
AGAGTCATAA GTAATTAAT CTGAATGTGT GTAGTTTAAT GGAATGGGA AAAAGATGGG 300
GGAAATGGAT GGAAGGTCTT GGACTCTGAG ACAAGGGGTC TATAATCAGT CCAITTTCAIT 360
ATTTCTAGCT TCCACCTTCA CCAAGGCAGA CAAGGAGGGC CCACCTCAGC TCTCTGCTC 420
65 CCCCTCCCTT TCCCACTAT TCATGTGTGC AAGAGTGCCC TGTCCACAG AACACGGGGA 480
ACAACCATCT CAATGACAAG GACAGCAGGT GGCAGGGCTC AACAGGACTC AGATGTCCCC 540
CCAGGGTTAA CTATGAAAC CCTCCATGAA GCCTGTCTGT CACCCCTCCC TCAAGGCAAG 600
CCCTGCACCT GGGTCTGAGG ATGAGGGTGG CAGTGAAAAT TAGGCCAGTG ACATCATTTT 660
CAGCCAGCTA GTGCCAAAAA ATATCAGGTG GTGTTTATCA AATAAGCCGA GCCAACCGGT 720
70 GATGAGGATG GTAGTGTGAG TCATGTGTGA CAGGTGAGGA ATGAAACAG AGTGCCCGAG 780
AGCTTCTATT TCCTTGAGGC AGGGCTCATT CATCTTATAA AAGCCAGCTG GCCATTGCTC 840
TCACACAAA CCCAAGGGAC CACACAGCCC ATTCTGTCC GTATACCAGG TAAGTCTCTG 900
ATTGCAACAA ACTGGCAATT CTAGTGTACT TTTTCATTAT TAGAAATTAG CTAAGGCCAA 960
ATATGTGTAA GCAGGTTAAT CCAGGGTTTC AATGGGAGAT AGAGAATAGT GGAATATCTT 1020
TATTTTAAGT TAAATTACAG TCTGGATTGT AAAGGACCTT AGAGATGGTT AGGGCTCCCA 1080
75 CCTCAGTAGA TAGTCATTGA ACTGGGATC CTGGAGAAGA TTGTTCAAAT GCCCATGGGA 1140
AGTTTCATAGC AGAAGTAGAA CTCAGGCCAG AGCACTCTCA GTAACACTGC AATTTCCCCC 1200
TGACAAGATA TTTATAGAAA TTTTAATTTA TTAGATGGAT CTCTACTGAG CATTTATTCC 1260
ATTTAAGGCA GTATGCTAGG CACTTTGGAC AAATCAATGC CCTAACGTAC TTACTTAACA 1320
AACAATAAAC CTAGCAGGAA GGTAAATACAT ATATATAAAT AARTGAAATG CAAAGTAGAT 1380
80 AGTAATTGCG ATGACGGAGA TGGGCAGAGA AGGGCTGTGC ACTTTTGGGA GACTTGCTCA 1440
AGGAGACCTC TAGGGTGTCA AGTGATGTGA GCTATGATGG AGGGGTATTG GGACAAGCAG 1500
AGATGGGAG AAAAGCATTT GGAAGGGACT GTGTAGCAC AGACCAGAAG CAAAACATA 1560
GAGGCTTAGA TGAATATAA GCCATCCTAT AAGTCACAGG CTTTCTACAT GGTACTAGGA 1620
GAGGAAAGTG GTCTGATGCC ATTTTCCAAA AGACCTAATA TCGGACCTC ATGTCCCTCA 1680
85 GAAGCCAGCT TTAGTAGGCG ATTTTCCAG AACAGATATA AGGTGCCTTG GGTAGGAAGG 1740
GAGCCAGAA GAGAAGCTCA ATAAAATGGA GCAGAAGAAA TTGCCTTTTA GCTCCTCTC 1800
TTCAAGGGC CTGAAAATTA TCCAAGCTTA TTTCAITTTT AAATGTAATG GGGGAGCTAA 1860

GGGAGATGAA AGGCTTTCTC TTCTAAAGGG TCCTGAAATA AAATCTGTTT GGCATTGAAT 1920
 TTGTATCCAT CTTTCTTTAA TTGAATCACT GTGTCAAGCTT TCTGTCTCTA GAAAAAACA 1980
 CATTTGAAGC ATGAATTCTC AGCAGCAGAA GCAGCCTTGC ACCCCACCCC CTCAGCCTCA 2040
 GCAGCAGCAG GTGAAACAAC CTTGCCAGCC TCCACCCAGG GAACCATGCA TCCCAAAAC 2100
 CAAGGAGCCC TGCCAAACCA AGGTGCCTGA GCCCTGCCAC CCAAAAGTGC CTGAGCCCTG 2160
 CCAGCCCAAG ATTCCAGAGC CCTGCCAGCC CAAGGTGCCT GAGCCCTGCC CTTCACCGT 2220
 CACTCCAGCA CACGCCAGC AGAAGACCAA GCAGAAGTAA TGTGGTCCAC AGCCATGCC 2280
 TTGAGGAGCT GGCCACTGGA TACTGAACAC CTTACTCCAT TCTGCTTATG AATCCCATTT 2340
 GCCTATTGAC CCTGCAGTTA GCATGCTGTC ACCCTGAATC ATAATCGCTC CTTTGCACTT 2400
 CTAAAAAGAT GTCCCTTACC CTCATTCTGG AGGCTCTCTG GCCTCTGCGT AAGGCTGAAC 2460
 GTCTCACTGA CTGAGCTAGT CTTCTTGTG CTCGGGTGCA TTTGAGGATG GATTGGGGGA 2520
 AGTCAAGTG ACCATCCCTA G

Seq ID NO: 51 Protein sequence:
 Protein Accession #: AAC26838

1 11 21 31 41 51
 | | | | |
 MNSQQQKQPC TFPFQPPQQQ VKQPCQPPPO EPCIPKTKEP CQPKVPEPCN PKVPEPCQPK 60
 IPEPCQKVP EPCPSTVTPA PAQKTKQK

Seq ID NO: 52 DNA sequence
 Nucleic Acid Accession #: NM_002638.1
 Coding sequence: 120-473

1 11 21 31 41 51
 | | | | |
 CAATACAGCT AAGGAATTAT CCCCTGTAAA TACCACAGAC CCGCCCTGGA GCCAGGCCAA 60
 GCTGACTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120
 TGAGGGCCAG CAGCTTCTTG ATCGTGTGGG TGTTCCTCAT CGCTGGGACG CTGTTCTTAG 180
 AGGCAGCTGT CAGCGGAGTT CCTGTTAAAG GTCAAGACAC TGTCAAAGCG CTGTTCCAT 240
 TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAAGTAA AGGTCAAGAT AAGTCAAAG 300
 CGCAAGAGCC AGTCAAAGGT CAGTCTCCA CTAAGCCTGG CTCCTGCCCC ATTATCTTGA 360
 TCCGGTGGCG CATGTTGAAT CCCCCTAACC SCTGCTTGA AGATACTGAC TGCCAGGAA 420
 TCAAGAAGTG CTGTGAAGGC TCTTGCGGGA TGGCCTGTTT CGTTCCCCAG TGAAGGGAGC 480
 OGGTCTTGC TGCACTGTG CGTCCCCAG AGCTACAGGC CCACTCTGCT CTTAAGTCCC 540
 TGCTGCCCTT CCCCCTCCCA CACTGTCCAT TCTTCTCTCC ATTCAAGATG CCCACGGCTG 600
 GAGCTGCCCT TCTCATCCAC TTTCCAATAA A

Seq ID NO: 53 Protein sequence:
 Protein Accession #: NP_002629.1

1 11 21 31 41 51
 | | | | |
 MRASSFLIVV VFLIAGTLVL EAAVTGVFVK QQDITVGRVP FNGQDPVKQ VSVKGQDKVK 60
 AQEPVKGPVS TKPGSCPILL IRCAMLNPPN RCLKDITDCPG IKKCCGSGCG MACFVPO

Seq ID NO: 54 DNA sequence
 Nucleic Acid Accession #: NM_019618
 Coding sequence: 75-584

1 11 21 31 41 51
 | | | | |
 GGCACGAGCC ACGATTCACT CCCCTGGACT GTAGATAAAG ACCCTTCTCT GCCAGGTGCT 60
 GAGACAACCA CACTATGAGA GGCACCTCAG GAGACGCTGA TGGTGGAGGA AGGGCCGCTC 120
 ATCAATCAAT GTGTAAACCT ATTACTGGGA CTATTAATGA TTTGAATCAG CAAGTGTGGA 180
 CCCTTCAGGG TCAGAACCTT GTGGCAGTTC CACGAAGTGA CAGTGTGACC CAGTCACTG 240
 TTGCTGTTAT CATATGCAAG TATCCAGAGG CTCTTGAGCA AGGCAGAGGG GATCCCATTT 300
 ATTTGGGAAT CCAGAAATCCA GAAATGTGTT TGTATTGTGA GAAGGTGGA GAACAGCCCA 360
 CATTGAGCTT AAAAGAGCAG AAGATCATGG ATCTGTATGG CCAACCCGAG CCGTGAAAC 420
 CCTTCTCTTT CTACCGTGCC AAGACTGGTA GGACCTCCAC CTTGAGTCT GTGGCCTTCC 480
 CGGACTGGTT CATTCCTCC TCCAAGAGAG ACCAGCCCAT CATTCGACT TCAGAACTTG 540
 GGAAGTCATA CAACACTGCC TTTGAATTAA ATATAAATGA CTGAATCAG CCTAGAGGTG 600
 GCAGCTTGGT CTTTGTCTTA AAGTTTCTGG TTCCCAATGT GTTTTGTCT ACATTTTCTT 660
 AGTGTCAATT TCACGCTGGT GCTGAGACAG GGGCAAGGCT GCTGTTATCA TCTCATTTTA 720
 TAATGAAGAA GAAGCAATTA CTTATAGCA ACTGAAGAAC AGGATGTGGC CTCAGAAGCA 780
 GGAGAGCTGG GTGGTATAAG GCTGTCTCT CAAGCTGGTG CTGTGTAGGC CACAAGGCAT 840
 CTGCATGAGT GACTTTAAGA CTCAAAGACC AAACACTGAG CTTTCTTCTA GGGGTGGGTA 900
 TGAAGATGCT TCAGAGCTCA TGGCGTTAC CCACGATGGC ATGACTAGCA CAGAGCTGAT 960
 CTCGTCTTCT GTTTTGTCTT ATCCCTCTT GGGATGATAT CATCCAGTCT TTATATGTTG 1020
 CCAATATACC TCATTGTGTG TAATAGAACC TTCTTAGCAT TAAGACCTTG TAAACAAAAA 1080
 TAATCTCTGT GTTAAGTTAA ATCATTTTG TCCTAATTGT AATGTGTAAT CTTAAGTTA 1140
 AATAAACTTT GTGTATTAT ATAATAAAAA AAAAAAAAAA AAA

Seq ID NO: 55 Protein sequence:
 Protein Accession #: NP_062564

1 11 21 31 41 51
 | | | | |
 MRGTPGDADG GGRAVQSMC KPITGTINDL NQVWTLQGG NLVAVPRSDS VTPVTVAVIT 60
 CKYPEALEQG RGDPIVLGIQ NPENCLYCEK VGEPTLQLK EQKIMDLVYG PEVPKPFIFY 120
 RAKTGTSTIL ESVAFFDWFII ASSKRQDPII LTSELGKSYN TAFELNIND

Seq ID NO: 56 DNA sequence
 Nucleic Acid Accession #: NM_003125
 Coding sequence: 65-334

1 11 21 31 41 51
 5 AGCAGTTCTA AGGGACCATA CAGAGTATTC CTCTCTTCAC ACCAGGACCA GCCACTGTTG 60
 CAGCATGAGT TCCAGCAGC AGAAGCAGCC CTGCATCCCA CCCCTCAGC TTCAGCAGCA 120
 GCAGGTGAAA CAGCCTTGCC AGCCTCCACC TCAGGAACCA TGCATCCCCA AAACCAAGGA 180
 GCCCTGCCAC CCAAGGTGC CTGAGCCCTG CCACCCCAAA GTGCCGTGAGC CCTGCCAGCC 240
 CAAGCTTCCA GAGCCATGCC ACCCAAGGT GCCTGAGCCC TGCCCTTCAA TAGTCACTCC 300
 10 AGCACAGCC CAGCAGAAGA CCAAGCAGAA GTAATGTGGT CCACAGCCAT GCCCTTGAGG 360
 AGCCGGCCAC CAGATCTGA ATCCCTATC CCATTCTGTG TATGAGTCCC ATTTGCTTGG 420
 CAATTAGCAT TCTGTCTCCC CCAAAAAAGA ATGTGCTATG AAGCTTTCTT TCCTACACAC 480
 TCTGAGTCTC TGAATGAAGC TGAAGGTCTT AGTACCAGAG CTAGTTTTC A GCTGCTCAGA 540
 ATTCACTGA AGAGAGACTT AAGATGAAAG CAAATGATTC AGCTCCCTTA TACCCCAT 600
 15 AAATTCACTT TCAATTCCA

Seq ID NO: 57 Protein sequence:

Protein Accession #: NP_003116

20 1 11 21 31 41 51
 MSSQQQKQPC IPPPQLQQQ VKQPCQPPPP EPCIPKTKEP CHPKVPEPCH PKVPEPCQPK 60
 25 LPEPCHPKVP EPCPSIVTPA PAQKTKQK

Seq ID NO: 58 DNA sequence

Nucleic Acid Accession #: NM_001793.2

Coding sequence: 71-2560

30 1 11 21 31 41 51
 AAAGGGGCAA GAGCTGAGCG GAACACCGGC CCGCGGTGCG GGCAGCTGCT TCACCCCTCT 60
 CTCTCAGGCC ATGGGGCTCC CTCTGTGACC TCTCGGTCTC CTCTCTCTTC TCAGGTTTGG 120
 CTGCTGTCAG TGCGCGGCTT CCGAGCGGTG CCGGGCGGTC TTCAGGGAGG CTGAAGTGAC 180
 35 CTGAGGAGCG GGAGGCGCGG AGCAGGAGCC CGGCCAGGCG CTGGGGAAG TATTCACTGG 240
 CTGCGCTGGG CAAGAGCCAG CTCTGTTTAA CACTGATAAT GATGACTTCA CTGTGCGGAA 300
 TGCGGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAAGG AATCCATTGA AGATCTTCCC 360
 ATCCAAACGT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCCAA TATCTGTCCC 420
 TGAATAATGG AAGGGTCCCT TCCCCAGAG ACTGAATCAG CTCAGTCTA ATAAAGATAG 480
 40 AGACACCAAG ATTTTCTACA GCATCAAGGG GCGGGGGGCA GACAGCCCCC CTGAGGGTGT 540
 CTTCGCTGTA GAGAAGGAGA CAGGCTGGTT GTTGTGAAT AAGCCACTGG ACCGGGAGGA 600
 GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTCAGAG AATGGTGCTT CAGTGGAGGA 660
 CCCCATGAAC ATCTCCATCA TCGTGACCGA CCAGAATGAC CACAAGCCCA AGTTTACCCA 720
 GGACACTTTC CGAGGGGCTC TCTTAGAGGG AGTCTACCA GGTACTTCTG TGAATGAGGT 780
 45 GACAGCCACG GATGAGGATG ATGCCATCTA CACCTACAAT GGGGTGGTGT CTACTCCAT 840
 CCATAGCCAA GAACCAAGG ACCCACAGGA CCTCATGTTT ACCATTACCC GGAGCACAGG 900
 CACCATCAGC GTCACTTCCA GTGGCCTGGA CCGGGAAAAA GTCCCTGAGT ACACACTGAC 960
 CATCCAGGCC ACAGACATGA ATGGGGAAGG CTCACACACC ACGGCAGTGG CAGTAGTGGA 1020
 GATCCCTGAT ATGCAATGCA ATGCTCCCAT GTTTGACCCC CAGAAGTACG AGGCCCATGT 1080
 50 GCCTGAGAAAT GCAGTGGGCC ATGAGGTGCA GAGGCTGACG GTCACTGATC TGGACGCCCC 1140
 CAACTCACCA GCGTGGCGTG CCACCTACCT TATCATGGGC GGTGACGACG GGGACCATTT 1200
 TACCATCACC ACCCACTCTG AGAGCAACCA GGCATCTCTG ACAACCCAGG AGGGTTTGGG 1260
 TTTTGAGGCC AAAAACCCAG ACACCTGTA CGTTGAAGTG ACCAACGAGG CCCCTTTTGT 1320
 GCTGAAGCTC CCAACTTCCA CAGCCACCAT AGTGGTCCAC GTGAGGATG TGAATGAGGC 1380
 55 ACCTGTGTTT GTCCCAACCT CCAAGTCTGT TGAGGTCCAG GAGGGCATCC CCACTGGGGA 1440
 GCCTGTGTGT GTCTACACTG CAGAAGACCC TGACAAGGAG AATCAAAAGA TCAGTACCG 1500
 CATCCTGAGA GACCAGCAGG GGTGGCTAGC CATGGACCCA GACAGTGGGC AGGTACAGC 1560
 TGTGGGACCC CTGACCGGTG AGGATGAGCA GTTTGTGAGG AACACATCT ATGAAGTCAAT 1620
 60 GGTCTTGGCC ATGACAAATG GAAGCCCTCC CACCACTGGC ACGGGAACCC TTCTGTCTAA 1680
 ACTGATTGAT GTCAATGACC ATGGCCCACT CCTGAGCCC CGTCAGATCA CCATCTGCAA 1740
 CCAAGCCCTT GTGCGCCAGG TGCTGAACAT CACGACAAAG GACCTGTCTC CCCACACCTC 1800
 CCTTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGGACGGCAG AGGTCAAAGA 1860
 GGAAGGTGAC ACAGTGTCTT TGTCCCTGAA GAAGTTCCTG AAGCAAGGATA CATATGACGT 1920
 65 GCACTTTTCT CTGTCTGACC ATGGCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980
 GTGCGACTGC CATGGCCATG TCGAAACCTG CCTTGGACCC TGAAGGGAGG GTTTCATCCT 2040
 CCCTGTGCTG GGGGCTGTCC TGGCTCTGCT GTTCTCTCTG CTGTGCTGCT TTTTGTGGT 2100
 GAGAAAGAAG CGGAAGATCA AGGAGCCCTT CCTACTCCCA GAAGATGACA CCCGTGACAA 2160
 CGTCTTCTAC TATGGCGAAG AGGGGGGTGG CGAAGAGGAC CAGGACTATG ACATCACCCA 2220
 GCTCCACCGA GGTCTGGAGG CCAGGCCGGA GGTGGTTCTC CGCAATGAGC TGGCAACCAAC 2280
 70 CATCATCCCG ACACCATGT ACCGTCTCTG GCCAGCCAAC CCAGATGAAA TCGGCAACTT 2340
 TATAATTGAG AACCTGAAGG CGGCTAACAC AGACCCACCA GCCCCGCCCT ACGACACCTT 2400
 CTTGGTGTTC GACTATGAGG GCAGCGGCTC CGACGCGCGT TCCCTGAGCT CCCTCACCTC 2460
 CTCGCGCTCC GACCAAGACC AAGATTACGA TTATCTGAAC GAGTGGGGCA GCCGCTTCAA 2520
 GAAGCTGGCA GACATGTACG GTGGCGGGGA GGAACGACTAG GCGGCTGCTC TGCAGGGCTG 2580
 75 GGSACCAAAC GTCAAGCCAC AGAGCATCTC CAAGGGGTCT CAGTTCCCCC TTCAGCTGAG 2640
 GACTTCGGAG CTGTCAAGGA AGTGGCCGTA GCAACTTGGC GGAGACAGGC TATGAGTCTG 2700
 ACGTTAGAGT GGTGTCTTCC TTAGCCTTTC AGGATGGAGG AATGTGGGCA GTTTGACTTC 2760
 AGCACTGAAA ACCCTCCACG CTGGGCCAGG GTTGCTCTAG AGGCCAAGTT TCCAGAGGCC 2820
 TCTTACTGCG GGTAAATGTC TCAACCTGT GTCTTGGGCC TGGGCTGCTG GTGACTGACC 2880
 80 TACAGTGGAC TTTCTCTCTG GAATGGAACC TTCTTAGGCC TCCTGGTGCA ACTTAATTTT 2940
 TTTTATTAAT GCTATCTTCA AAACGTTAGA GAAAGTCTT CAAAAGTGCA GCCCAGAGCT 3000
 GCTGGGCCCA CTGGCCCTCC TGCATTTCTG GTTTCAGAC CCAATGCCT CCCATTGGGA 3060
 TGGATCTCTG CGTTTATATA CTGAGTGTGC CTAGTGTGCC CCTTATTTT TATTTTCCCT 3120
 85 GTTGGCTTGC TATAGATGAA GGGTGAGGAC AATCGTGTAT ATGTACTAGA ACTTTTAT 3180
 TAAAGAACT TTTCCAGAA AAAA

Seq ID NO: 59 Protein sequence:

Protein Accession #: NP_001784.2

1 11 21 31 41 51
5 MGLRGRPLAS LLLIQVCWLQ CAASEPCRAV FREAEVTLEA GGAEQEPGQA LGRVFMGCPG 60
QEPALFSTDN DDFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG 120
KGPFPQRLNQ LKSNKORDTK IFYSITGPGA DSPPEGVFAV EKETGMLLN KPLDREIAR 180
YELFGHAVSE NGASVEDPMN ISIIIVTDQND HKPKPTQDTF RGSVLEGVLF GTSVMQVTAT 240
DEDDAIYTYN GVVAYSIHSQ EPKDPHDLME TIHRSTGTIS VISSGLDREK VPEYTLTIQA 300
10 TDMDGDGSTT TAVAVVEILD ANDNAPMFDE QKYEAVHPEN AVGHEVQRLT VTDLDAPNSP 360
AWRATYILIM GDDGDHFTIT THPESNQIL TTRKGLDFAE KNQHTLYVEV TNEAPFVLKL 420
PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGIPTGEFVC VYTAEDPDKE NQKISYRILR 480
DPAGWLAMPD DSGQVTAVGT LDREDEQFVR NNIYEVMLA MDNGSPPTTG TGTLLLTLD 540
VNDHGPVPEP RQITICNQSP VRQVLNITDK DLSPTSPFQ AQLTDDSDIY WTAEVNNEGD 600
15 TVVLSLKKFL KQDTYDVHLS LSDHGNKEQL TVIRATVDCD HGHVETCPGP WKGGFLLPVL 660
GAVLALLFLL LVLLLVLRKK RKIKEPLLLP EDDTRDNVFI YGEEGGGEED QDVDTQLHR 720
GLEARPEVVL RNDVAPITIP TPMYRPRPAN PDEIGNFIIE NLKAAANTDPT APPYDTLLVF 780
DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD

20 Seq ID NO: 60 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 162-428

1 11 21 31 41 51
25 GCGTTCGGTT GGCGGCGGAT TCGAACCTTC GGAAGTGGGT TTTTCTGCCT GAAGAAGCGT 60
CATACCGACC GGATGTGTTT CGCTGGCCCA GTGTCCCGG AGCTTGTGTG CGATACAGAG 120
AGCACTCCGG AAGCTGAGGC AGCTGGTACT TGACAGAGAG GATGGCGCTG TCGACCATAG 180
TCTCCAGAG GAAGCAGATA AAGCGGAAG CTCCCGTGG CTCTCTAAAG CGAGTCTTCA 240
30 AGCGAAGAA GCCTCAACTT CGTCTGGAGA AAAGTGGTGA CTTATTGGTC CATCTGAAT 300
GTTTACTGTT TGTTCATCGA TTAGCAGAAG AGTCCAGGAC AAACGCTTGT GCGAGTAAAT 360
GTAGAGTCAT TAACAAGGAG CATGTACTGG CCGCAGCAAA GGTAATTCTA AAGAAGAGCA 420
GAGGTAGAA GTCAAGAAGC ATATTCTTGA AAGTTATGAT GCATTCTTTT GGGTGGTAAC 480
AGATCATAAA GACATTTTTT ACACATCAAT TAATATGGGA TTATTAAATA TTGG

35 Seq ID NO: 61 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
40 MALSTIVSQR QKIKRKAPRG FLKRVFKRKK PQLRLEKSGD LLVHLNCLLF VHLRAESRT 60
NACASKRVI NKEHVLAAAK VILKKSRG

45 Seq ID NO: 62 DNA sequence
Nucleic Acid Accession #: NM_000094.2
Coding sequence: 99-8933

1 11 21 31 41 51
50 GGGCTGGAGG GGGCGTGGGC TCGGACCTGC CAAGGCCACC GCAGGGGGGA GCAAGGGACA 60
GAGGCGGGGG TCCTAGCTGA CGGCTTTTAC TGCTAGGAT GAGCGTGGG CTTCTGGTGG 120
CGCGCTCTG CGCGGGGATC CTGGCAGAGG CGCCCGAGT GCGAGCCAG CACAGGGAGA 180
GAGTGACCTG CACGCGCTTT TACGCGCTG ACATTGTGTT CTTACTGGAT GGCTCCTCAT 240
CCATTGGCGG CAGCAATTTT CGCGAGGTCC GCAGCTTTCT CGAAGGGCTG GTGTGCTT 300
55 TCTCTGGAGC AGCGAGTGA CAGGGTGTGC GCTTTGCCAC AGTGCAGTAC AGCGATGACC 360
CACGGACAGA GTTGGCGCTG GATGCACTTG GCTCTGGGGG TGATGTGATC CGCGCCATCC 420
GTGAGCTTAG CTACAAGGGG GGCAACACTC GCACAGGGGC TGCAATTCTC CATGTGGCTG 480
ACCATGTCTT CCGTCCCGAG CTGGCCCGAC CTGGTGTCCC CAAGGTCTGC ATCTGTATCA 540
60 CAGACGGGAA GTCCCGAGAC CTGGTGGACA CAGCTGCCCA AAGGCTGAAG GGGCAGGGGG 600
TCAAGCTATT TGCTGTGGGG ATCAAGAATG CTGACCCCTGA GGAGCTGAAG CGAGTTGCC 660
CACAGCCAAC CTCGACTTTC TTCTTCTTGG TCAATGACTT CAGCATCTTG AGGACACTAC 720
TGCCCTCTGT TTCCCGGAGA GTGTGCACGA CTGCTGGTGG GTGCGCTGTG ACCCGACCTC 780
CGGATGACTC GACCTCTGCT CCAAGAGACC TGCTGTGCTG TGAGCCAAGC AGCCAACTCT 840
TGAGAGTACA GTGGACAGCG GCCAGTGGCC CTGTGACTGG CTACAAGGTC CAGTACACTC 900
65 CTCTGACGGG GCTGGGACAG CCACTGCCGA GTGAGCGGCA GGAGGTGAAC GTCCAGCTG 960
GTGAGACCAG TGTGGCGCTG CGGGGTCTCC GGCCACTGAC CGAGTACCAA GTGACTGTGA 1020
TTGCCCTCTA CGCCCAACAG ATCGGGGAGG CTGTGAGCGG GACAGCTCGG ACCACTGCC 1080
TAGAAGGGCC GGAAGTGACC ATCCAGAATA CCACAGCCCA CAGCCTCCTG GTGGCTTGGC 1140
GGAGTGTGCC AGGTGCCACT GGCTACCGTG TGACATGGCG GGTCTCTCAG GTTGGGCCCA 1200
70 CACAGACGCA GGAGCTGGGC GTGGGACCG GTTCAGTGTG GCTGCGTGAC TTGGAGCCTG 1260
GCACGGACTA TGAGGTGACC GTGAGCACCC TATTGGCCCG CAGTGTGGGG CCGGCCACTT 1320
CCCTGATGGC TCGCACTGAC GCTTCTGTGG AGCAGACCCCT GCGCCCGGTC ATCCTGGGCC 1380
CCACATCCAT CCTCTTTTCC TGGAACTTGG TGCTGAGGC CCGTGGCTAC CGTGTGGAAT 1440
GGCGGGGTGA GACTGGCTTG GAGCCACCGC AGAAGGTGGT ACTGCCCTCT GATGTGACCC 1500
75 GCTACCACTA GGATGGGCTG CAGCGGGGCA CTGAGTACCG CCTCACACTC TACACTCTGC 1560
TGGAGGGCCA CGAGTGGGCC ACCCTGCAA CCGTGGTTC CACTGGACCA GAGCTGCGCT 1620
TGAGCCCTGT AACAGACCTG CAAGCCACCG AGCTGCCCGG GCAGCGGGTG CAGTGTCTCT 1680
GGAGCCGAGT CCTGGTGGCC ACCCAGTACC GCATCATTGT CTTGGATGAC GTTCAGGCTG 1740
AGCGGACCTT GGTGCTTCCT GGGAGTCAGA CAGCATTCGA CTTGGATGAC GTTCAGGCTG 1800
80 GCTTAGCTG GATGCTGGG GTGTCTGCTC GAGTGGGTCC CCGTGGGGGC AGTCCAGTGC 1860
TCCTCACTGT CCGCCGGGAG CCGGAACTC CACTTGTCTG TCCAGGGCTG CGGCTGTGG 1920
TGTGATGTC AACCGAGTG AGGGTGGCCT GGGGACCCGT CCTTGGAGCC AGTGGATTTC 1980
GGATTAGCTG GAGCACAGGC AGTGGTCCGG AGTCCAGCCA GACACTGCCC CCAGACTCTA 2040
CTGCCACAGA CATCACAGGC CTGCAGCCTG GAACCACTA CAGGTGGGCT GTGTGGGTAC 2100
85 TGGAGGCGAG AGAGGAGGGC CCTGCTGAG TCATCGTGGC TCGAACGGAC CCACTGGGGC 2160
CAGTGAGGAC GGTCCATGTG ACTCAGGCCA GCAGCTCATC TGTCAACATT ACCTGGACCA 2220
GGGTCTCTGG CGCCACAGGA TACAGGGTTT CCGTGGCACT AGCCCAACGG CAGAGAAAT 2280

	CCCAGTTGGT	TTCTGGGGAG	GCCACGGTGG	CTGAGCTGGA	TGGACTGGAG	CCAGATACTG	2340
	AGTATACGGT	GCATGTGAGG	GCCCATGTGG	CTGGCGTGGA	TGGGCCCCCT	GCCTCTGTGG	2400
	TTGTGAGGAC	TGCCCTCTAG	CCTGTGGGTC	GTGTGTGAGG	GCTGCAGATC	CTCAATGCTT	2460
5	CCAGGACGCT	TCTACGGATC	ACCTGGGTAG	GGGTCACTGG	AGCCACAGCT	TACAGACTGG	2520
	CCTGGGGCCG	GAGTGAAGCG	GGCCCCATGA	GGCACAGAT	ACTCCCAGGA	AACACAGACT	2580
	CTGCAGAGAT	CCGGGGTCTC	GAAGGTGGAG	TCAGCTACTC	AGTGCAGATG	ACTGCACCTG	2640
	TCGGGGACCG	CGAGGGCACA	CCTGTCTCCA	TTGTGTTCAC	TACGCCGCTT	GAGGCTCCGC	2700
	CAGCCCTGGG	GACGCTTCAC	GTGGTGCAGC	GCGGGGAGCA	CTGCTGAGG	CTGCGCTGGG	2760
	AGCCGGTGCC	CAGAGGCGAG	GGCTTCCTTC	TGCACTGGCA	ACCTGAGGGT	GGCCAGGAAC	2820
10	AGTCCCGGGT	CCTGGGGCCC	GAGCTCAGCA	GCTATCACCT	GGACGGGCTG	GAGCCAGCGA	2880
	CACAGTACCG	CGTGAGGCTG	AGTGTCTTAG	GGCCGGCTGG	AGAAGGGCCC	TCTGCAGAGG	2940
	TGACTGCGCG	CACGTAGTCA	CCTCGTGTTC	CAAGCATTGA	ACTAGTGTGT	GTGGACACCT	3000
	CGATCGACTC	GGTGACTTTG	GCCTGGACTC	CAGTGTCCAG	GGCATCCAGC	TACATCTCAT	3060
	CCTGGGGGCG	ACTCAGAGGC	CCTGGCCAGG	AAGTGCCTGG	GTCCCGCAG	ACACTTCCAG	3120
15	GGATCTCAAG	CTCCACGCGG	GTGACAGGGC	TAGAGCCTGG	CGTCTCTTAC	ATCTTCTCCC	3180
	TGACGCGCTGT	CCTGGATGCT	GTGGGGGCTC	CTGAGGCATC	TGTACACAG	ACGCCAGTGT	3240
	GGCCCGCTGG	TCGGGGCGAT	GTGGTGTTC	TACCAATGCG	CACTCAAGAC	AATGCTCACC	3300
	GTGGGAGGCG	TACGAGGAGG	GTCTGGAGC	GTCTGTGTTT	GGCACTTGGG	CCTCTTGGGC	3360
	CACAGGCAGT	TCAGGTTGGC	CTGCTGTCTT	ACAGTCATCG	GCCTTCCCA	CTGTCTCCAC	3420
20	TGAATGCGTC	CCATGACCTT	GGCATTATCT	TGCAAGGAT	CCGTGACATG	CCCTACATGG	3480
	ACCCAGTGGT	GAACAACCTG	GGCACAGCCG	TGGTCAAGC	TCACAGATAC	ATGTTGGCAC	3540
	CAGATGCTCG	TGGCGGCGCG	CAGCACTGAC	CAGGGGTGAT	GGTCTGCTA	GTGGATGAAC	3600
	CCTTGAGAGG	TGACATATTC	AGCCCCATCC	GTGAGGCCCA	GGCTTCTGGG	CTTAATGTGG	3660
	TGATGTTGGG	AATGGCTGGA	GCGGACCCAG	AGCAGCTGCG	TCGCTTGGCG	CCGGGTATGG	3720
25	ACTCTGTCCA	GACCTTCTTC	GCCGTGGATG	ATGGGCCAAG	CCTGGACACG	GCAGTCAGTG	3780
	GTCTGGCCAC	ATGCCCTGTGT	CAGGCATCCT	TCACACTCA	GCCTCGGCCA	GAGCCCTGCC	3840
	CAGTGTATGG	TCCAAAGGCG	CAGAAGGGGG	AACCTGGAGA	GATGGGCTGG	AGAGGACAGG	3900
	TTGGGCTCTC	TGGCGACCTT	GGCCTCCCGG	GCAGGACCGG	TGCTCCCGGC	CCCCAGGGGC	3960
	CCCCGTGAAG	TGCCACTGCC	AAGGGCGAGA	GGGGCTTCCC	TGGAGCAGAT	GGGCGTCCAG	4020
30	GCAGCCCTGG	CCGCGCCGGG	AATCCTGGGA	CCCCGTGAGC	CCCTGGCCTA	AAGGGCTCTC	4080
	CAGGGTTCCT	TGGCCCTCGT	GGGGACCCGG	GAGAGCGAGG	ACCTCGAGGC	CCAAAGGGGG	4140
	AGCCGGGGCG	TCCCGGACAA	GTGATCGGAG	GTGAAGGACC	TGGGCTTCTC	GGGGGAAAG	4200
	GGGACCCCTG	ACCATCGGGC	CCCCCTGGAC	CTCGTGGACC	ACTGGGGGAC	CCAGGACCCC	4260
	GTGGCCCCCG	AGGGCTTCTT	GGAAACGCCA	TGAAGGTGA	CAAAGGCGAT	CGTGGGGAGC	4320
35	GGGGTCCCCC	TGGACAGGCT	GAAGGTGGCA	TGCTCTCTGG	GGAGCTTGGG	CTGCGGGGTC	4380
	TTCCCGGAAG	CCCTGACACC	CAAGGCCCGG	TTGGCCCCCG	TGGAAGAGAA	GGAGAAAAAG	4440
	GTGACTTTCG	GGATGGAGCT	CCAGGCTCTC	CAGGACAAAC	TGGGTCTCCG	GGTGAGCAGC	4500
	GCCACCGGGG	ACCTCTCTGA	GCTATTGGCC	CCAAAGGTGA	CCGGGGCTTT	CCAGGGCCCC	4560
	TGGGTGAGCG	TGGAGAGAA	GGCGAAAGTG	GACCCCCAGG	CCGAGCGGGA	TCCCGGGGCG	4620
40	TGCCAGGGGT	TGCTGGACGT	CCTGGAGCCA	AGGGTCTCTG	AGGGCCACCA	GGACCCACTG	4680
	GCGCGCAAGG	AGAGAAAGGG	GAGCCTGGTC	GCCCTGGGGA	CCCTGCACTG	GTGGGACCTG	4740
	CTGTGTCTGG	ACCCAAAGGA	GAAAAGGGAG	ATGTGGGGCC	CGCTGGGCCC	AGAGGAGCTA	4800
	CCGGAGTCCA	AGGGGAACCG	GGCCCCACCG	GCTTGGTTCT	TCCTGGAGAC	CCTGGCCCCA	4860
	AGGGAGACCC	TGGAGACCGG	GGTCCCATTG	GCCTTACTGG	CAGAGCAGGA	CCCCAGGTGG	4920
45	ACTCAGGGCC	TCTTGGAGAG	AAGGGAGACC	CTGGGCGGCC	TGGCCCCCA	GGACCTGTTG	4980
	GCCCGCGAGG	ACGAGATGCT	GAAGTTGGAG	AGAAAGGTGA	CGAGGGTCTT	CCGGGTGACC	5040
	CGGGTTTGGC	TGGAAAGGCA	GGCGAGCGTG	GCCTTCGGGG	GGCACCTGGA	GTTCGGGGCG	5100
	CTGTGGGTGA	AAAGGGAGAC	CAGGGAGATC	CTGGAGAGGA	TGACGAAAT	GGGACCCCTG	5160
50	GATCATCTGG	ACCCAAAGGT	GACCGTGGGG	AGCCGGGTCC	CCGAGGACCC	CCGGGACGGC	5220
	TGGTAGACAC	AGGACCTGGA	GCCAGAGAGA	AGGGAGAGCC	TGGGGACCGC	GGACAAGAGG	5280
	GTCTCTGAGG	GCCCAAGGCT	GATCCTGGCC	TCCCTGGAGC	CCCTGGGGAA	AGGGGCACTG	5340
	AAGGGTTTCG	GGGACCCGCA	GGGCCACAGG	GGGACCCAGG	TGTCGGAGGC	CACGACGGAG	5400
	AAAGGGGTGA	CCGGGGTCCC	CCTGGGCTGG	ATGGCCGGAG	CGGACTGGAT	GGGAAACCAAG	5460
	GAGCGCTGG	GGCCTCTGGG	CCGAATGGTG	CTGCAGGCA	AGCTGGGGAC	CCAGGGAGAG	5520
55	ACGGGCTTCC	AGGGCTTCTG	GGAGAACAG	GCCTCCTTGG	CCCTCTGGT	CCCCCTGGAT	5580
	TACCGGGAAA	GCCAGCGGAG	GATGGGAAAC	CTGGGCTGAA	TGGAAAAAAC	GGAGAACCTG	5640
	GGGACCTTGG	AGAGAGCGGG	AGGAAGGGAG	AGAAAGGAGA	TTCAGGCGCC	CTCGGGAGAG	5700
	AAGGTGCTGA	TGGCCCCAAG	GGTGAGCGTG	GAGCTCTTGG	TATCCTTGA	CCCCAGGGGC	5760
	CTCCAGGCTT	CCCAAGGCCA	GTGGGCCCTC	CTGGCCAGGG	TTTTCTTGGT	GTCCAGGGAG	5820
60	GCACGGGCCC	CAAGGGTGAC	CGTGGGGAGA	CTGGATCCAA	AGGGGAGCAG	GGCCTCCCTG	5880
	GAGAGCGTGG	CCTGCGAGGA	GAGCCTGGAA	GTGTGCCGAA	TGTGGATCGG	TGTCTGGAAA	5940
	CTGCTGGCAT	CAAGGCATCT	GCCCTGCGGG	AGATCGTGG	GACCTGGGAT	GAGAGCTCTG	6000
	GTAGCTTCTT	GCCTGTGCCC	GAACGGCGTC	GAGGCCCCAA	GGGGGACTCA	GGCGAACAGG	6060
	GCCCCCAGG	CAAGGAGGGC	CCCATCGGCT	TTCCTGGAGA	ACGCGGGCTG	AAGGGCGACC	6120
65	GTGGAGACCC	TGGCCCTCAG	GGGCCACCTG	GTCTGGCCCT	TGGGGAGAGG	GGCCCCCCCC	6180
	GGCCTTCCGG	CCTTGCCTGG	GAGCCTGGAA	AGCCTGGTAT	TCCCGGGCTC	CCAGGCAGGG	6240
	CTGGGGTGGT	GGGAGAGGCA	GGAAAGGCCA	GAGAGAGGGG	AGAAAGGGGA	GAGAAAGGAG	6300
	AACGTGGAGA	ACAGGGCAGA	GATGGCCCTC	CTGGACTCCC	TGGAACCCCT	GGGCCCCCGG	6360
	GACCCCTGG	CCCCAAGGTG	TCTGTGGATG	AGCCAGGTCC	TGGAATCTCT	GGAGAACAGG	6420
70	GACCCCTGG	ACTCAAGGGT	GCTAAGGGGG	AGCCGGGCG	CAATGGTGAC	CAAGGTCCCA	6480
	AAGGAGCAG	GGGTGTGCCA	GGCATCAAAG	GAGACCGGGG	AGAGCTTGA	CCGAGGGGTC	6540
	AGGACGGCAA	CCCGGGTCTA	CCAGGAGAGC	GTGGTATGGC	TGGGCTGAA	GGGAAGCCGG	6600
	GTCTGCGAGG	TCCAGAGGCG	CCCCCTGGCC	CAGTGGGTGG	TCACTGGAGC	CCTGGACCAC	6660
	CTGGTGGCCC	GGGTCTTGGT	GGCCCTGACG	GACCCCAAGG	ACCTTCTGGC	CTGAAGGGGG	6720
75	AGCCTGGAGA	GACAGGACCT	CCAGGACGGG	GCCTGACTGG	ACCTACTGGA	GCTGTGGGAC	6780
	TTCTTGGACC	CCCCGGCCCT	TCAGGCCCTT	TGGGTCCACA	GGGGTCTCCA	GGTTTGGCTG	6840
	GACAGTGGGG	GGAGACAGGG	AAGCCGGGAG	CCCCAGGTGG	AGATGGTGGC	AGTGGGAAAG	6900
	ATGGAGCAG	AGGGAGCCCT	GGTGTGCCAG	GGTCACAGG	TCTGCTTGGC	CCTGTGGGAC	6960
	CTAAAGGAGA	ACCTGGCCCC	ACGGGGGGCC	CTGACAGGC	TGTGGTGGGG	CTCCCTGGAG	7020
80	CAAGAGGAGA	GAAGGGAGCC	CCTGGAGGCC	TGCTGGAGA	CCTGTGGGGT	GAGCCGGGAG	7080
	CCAAAGGTGA	CCGAGGACTG	CCAGGGCCGC	GAGGCGAGAA	GGGTGAAGCT	GGCCGTGCAG	7140
	GGGAGCGGG	AGACCTTGGG	GAAGATGGTC	AGAAAGGGGC	TCCAGGACCC	AAAGGTTTCA	7200
	AGGGTGACCC	AGGAGTCTGG	GTCCCGGGCT	CCCCTGGGCC	TCCTGGCCCT	CCAGGTGTGA	7260
	AGGGAGATCT	GGGCTTCCCT	GGCCTGCCCC	GTGCTCTTGG	TGTTGTGGGG	TTCCCGGGTC	7320
85	AGACAGGGCC	TCGAGGAGAG	ATGGGTGAGC	CAGGCCCTAG	TGGAGAGCGG	GCTCTGGCAG	7380
	GCCCCCAGG	GAGAGAGGGA	ATCCAGGAC	CCCTGGGGCC	ACCTGGACCA	CCGGGGTTCAG	7440
	TGGGACACC	TGGGGCTCTT	GGACTCAAAG	GAGACAGGGG	AGACCTTGA	GTAGGGCTGC	7500

WO 02/086443

PCT/US02/12476

5 CTGGGCCCC AGGCGAGCGT GGGGAGCCAG GCATCCGGGG TGAAGATGCG CGCCCCGGCC 7560
 AGGAGGGGACC CCGAGGACTC ACGGGGCCCC CTGGCAGCAG GGGAGAGCGT GGGGAGAGAG 7620
 GTGATGTTTG GAGTGCAGGA CTAAGGGGTG ACAAGGGAGA CTCAGCTGTG ATCTGGGGG 7680
 CTOCAGGCCC ACGGGGTGCC AAGGGGGACA TGGGTGAACG AGGGCTCGG GCGTTGGATG 7740
 GTGACAAAGG ACCTCGGGGA GACATGGGG ACCCTGGTGA CAAGGGCAGC AAGGGAGAGC 7800
 CTGGTGACAA GGGCTCAGCC GGGTTGCCAG GACTGCGTGG ACTCCTGGA CCCCAGGGTC 7860
 AACCTGGTGC AGCAGGGATC CCTGGTGACC CGGGATCCCC AGGAAAGGAT GGAGTGCCCTG 7920
 GTATCCGAGG AGAAAAAGGA GATGTTGGCT TCATGGGTCC CCGGGGCTC AAGGGTGAAC 7980
 10 GGGAGGTGAA GGGAGCGCTG GGCCTTGATG GAGAGAAGGG AGACAAGGGA GAAGCTGGTC 8040
 CCCCAGGCCG CCCCAGGCTG GCAGGACACA AAGGAGAGAT GGGGGAGCCT GGTGTGCCG 8100
 GCCAGTCGGG GGGCCCTGGC AAGGAGGGCC TGATCGGTCC CAAGGCTGAC CGAGGCTTTG 8160
 ACGGCGAGCC AGGCCCCAAG GGTGACCAGG GCGAGAAAGG GGAGCGGGGA ACCCCAGGAA 8220
 TTGGGGGCTT CCGAGGCCCC AGTGGAATG ATGGCTCTGC TGGTCCCCCA GGGCCACCTG 8280
 15 GCGAGTGTTC CCGCAGAGC CCGGAAGGAC TTCAGGGCCA GAAGGGTGAG CGAGGTCCCC 8340
 CCGGAGAGAG AGTGGTGGGG GCTCCTGGGG TCCCTGGAGC TCCTGGCGAG AGAGGGGAGC 8400
 AGGGGCGGCC AGGGCCTGGC GGTCTCGAG GCGAGAAGGG AGAAGCTGCA CTGACGGAGG 8460
 ATGACATCCG GGGCTTTGTG CGCCAAGAGA TGAGTCAGCA CTGTGCCTGC CAGGGCCAGT 8520
 TCATCGCATC TGGATCAAGA CCCCCTCCCTA GTTATGCTGC AGACACTGCC GGCTCCACAG 8580
 20 TCCATGCTGT GCGTGTGCTC GCGCTCTCTC ATGCAGAGGA GGAAGAGCGG GTACCCCTG 8640
 AGGATGATGA TACTCTGAA TACTCCGAGT ATCTGTGGA GGAGTACGAG GACCTGAAC 8700
 CTCTTGGGA TAGTGATGAC CCGTGTTCCT TGCCACTGGA TGAGGCTCTC TGCACTGCC 8760
 ACACCTCGTG CTGGTACCAT CCGGCTGTGA CAGGCGAGC AGAGGCTGT CACCTTTTG 8820
 TCTATGGTGG CTGTGGAGGG AATGCCAACC GTTTTGGGAC CCGTAGGGCC TGGGAGCGCC 8880
 25 GCTGCCACCC CCGGGTGGTC CAGAGCCAGG GGAAGGTATC TGCCAGGAGC TGAGGCGCCAG 8940
 ATAATGAGCT GAGATTGAGC ATCCCTGGA GAGTCCGGGG TCTCAGCAGA ACCCACTGT 9000
 CCTCCCTCTT GGTGCTAGAG GCTTGTGTGC ACGTGAAGCT GCGAGTGCAC GTCCGTTATT 9060
 TCACTGACTT GGTCCCGTGG GTCTAGCCTT CCCCCCTGTG GACAAACCCC CATTTGGGCT 9120
 CCTGCCACCC TGGCAGATGA CTCACTGTGG GGGGGTGGCT GTGGGCAGTG AGCGGATGTG 9180
 30 ACTGGCGTCT GACCCGCCCC TTGACCCAAG CCGTGATGA CATGGTGTG ATTCGGGGG 9240
 GCATTAAAGC TGCTGTTTAA AAGGCCAAA AA

Seq ID NO: 63 Protein sequence:
 Protein Accession #: NP_000085.1

35 1 11 21 31 41 51
 MTLRLVAAL CAGILAEAPR VRAQHRERV CTRLYAADIV FLDDGSSSIG RSNFREVRSP 60
 LEGVLVPPSG AASAGQVRFA TVQYSDPRT EFGLDALGSG GDVIRAIREL SYKGNTRTG 120
 40 AAILHVADHV FLPLARPGV PKVCILITDG KSQDLVDTAA QRLKGGGVKL FAVGIKNADP 180
 EELKRVASQP TSDFFFFVND FSILRTLLEL VSRVCTTAG GVPVTRPDD STSAPRDLVL 240
 SEFSSQSLRV QMTAASGPVT GYKVQYTPLT GLGQPLPSER QEVNVFAGET SVRLRGLRPL 300
 TEYQVTVIAL YANSIGEAVS GTARTALEG PELTIQNTTA HSLLVANRSV PGATGYRVTW 360
 RVLSGGTQPT QELSGPGQSV LLRLDLEPGTD YEVTVSTLFG RSVGPATSLM ARTDASVEQT 420
 45 LRPVILGPTS ILLSNLVPE ARGYRLEWRN ETGLEPPQKV VLPSDVTRYQ LDGLQPGTEY 480
 RLTLTYLLEG HEVATPATVV PTGPPELVSP VTDLQATELP GQRVRSWNSP VPGATQYRII 540
 VRSTQGVERT LVLPSSQTAF DLDDVQAGLS YTVRVARSVG PREGSASVLT VRREPETPLA 600
 VPGLRVVSD ATRVRVWAMP VPGASGFRIS WSTGSGPSS QTLPPDSTAT DITGLQPGTT 660
 YQVAVSVLRG REEGPAIVIV ARTDPLGVR TVHVTQASS SVTITWTRVP GATGYRVSWH 720
 50 SAHGPEKSQL VSGEATVAEL DGLEPDTEYT VEVRAHVAGV DGPASVVRV TAPEPVGRVS 780
 RLQILNASSD VLRIWVGVT GATAYRLAWG RSEGGPMRHQ ILPGNTDSAE IRGLEGGVSY 840
 SVRTALVGD REGTPVSIIV TTPPEAPPAL GTLHVVRGE HSLRLRVEPV PRAQGLLHW 900
 QPEGQEQSR VLGPESLSYH LDGLEPATQY RVRLSVLGPA GEGPSAEVTA RTESPRVPSI 960
 ELRVVDTSID SVTLAWTPVS RASSYILSWR PLRGPQGEVP GSPQTLPGIS SSQRTVLEP 1020
 55 GVSIFYSLTP VLDGVRGPEA SVTQTPVCPR GLADVFLPH ATQDNHRAE ATRVRLERLV 1080
 LALGLPLGPA VQVGLLSYSH RPSPLFLPLNG SHDLGILQR IRDMPYMDPS GNNLTAVVT 1140
 AHRVYLPADA RGRQHVPGV MVLLVDEPLR GDIFSPIREA QASGLNVML MAGADPEQL 1200
 RRLAPGMDSV QTFFAVDGSP SLDQAVSGLA TALCQASFTT QPRPEPCFVY CPKQKQKGP 1260
 EMGLRGQVGP PGDPLGPGRT GAPGPQGGP SATAKGERGF PGADGRPGSP GRAGNPPTGP 1320
 60 APGLKSPPL PGPRGPGERG GPRGPKEGPG APQVIGGEG PGLPCRKGDG GSPGPPGPRG 1380
 PLGDPGPRGP PGLPGTAMKG DRGDRGERGP PGPGEGGIAP GEPGLPGLPG SPGPQGPVGP 1440
 PGKKKEGGS EDGAPGLPGQ PGSPGEQGP GPPGAIGPKG DRGFPGLGE AGEKGERGPP 1500
 GPAGSRGLPG VAGRPAKGP EGFPGPGRQ GEKGEPRGP DPAVVGPVA GPKEKGDVG 1560
 PAGPRGATGV QGERGPPGLV LPDGPMPKGD PGDRGPGLT GRAGPPGDSG PPEKGDGPR 1620
 65 PGPPGPVGR GRDGEVGEK DEGPDPDPGL PKKAGERGLR GAPGVRGPVG EKGDQGDGPE 1680
 DGRNGSPGSS GPKGDRGEPG PPGPPGRLVD TGPAREKEGE PGDRGQGGPR GPKGDPGLPG 1740
 APGERGIEG RGPQPGQDP GVRGPAGEKG DRGPPGLDGR SGLDGKPGAA GSPGPNGAAG 1800
 KAGDPGRDGL PGLRGEQGLP GSPGPPGLPG KPGEDGKPLG NGKNGEPGD GEDGRKGERG 1860
 DSGASGREGR DGPKGERGAP GILGPQGGP LPGPVGPVGG GPPVGPVGGT PKGDRGETGS 1920
 70 KGEQQLPGER GLRGEPSV NVDRLLLETAG IKASALREIV ETWDESSGSF LPVPERRRGP 1980
 KGDSEGGPP GKEGPIGFP ERGLKGRGD PGPQGPPLA LGERGPPGPS GLAGEPGKPG 2040
 IPGLPRAGG VGEAGRGER GERGEKGERG EQGRDGPPL PGTGPPGPP GPKVSVDEPG 2100
 PGLSGEGPP GLKAKGEPG SNGDQGPKG RGVPGIKGDR GEPGPRQDQ NPGLPGERGM 2160
 AGPEGKGLQ GPRGPPGPV GHGDPGPPGA PGLAGPAGP GPSGLKGEPE ETGPPGRGLT 2220
 75 GPTGAVLPG PPGPSGLVGP QSPGLPGQV GETGKPGAP RDGASGKGDG RSGPVPSP 2280
 GLPGVPGPK EPGPTGAPQ AVVGLPGAKG EKGAPEGLAG DLVGEPGAKG DRGLPGPRGE 2340
 KGEAGRAGEP RGPGEQKQK APGPKGFKD PGVGVPGSP PPGPPGVKGD LGPLPGAP 2400
 GVVGPPQQTG PRGEMQPGP SGERGLAGP GREGIPGLQ PPGPPGSPV PGASGLKGD 2460
 GDFGVGLPG RGERGEPGIR GEDGRPGQG PRGLTGPPTS RGERGEKGDV GSAGLKGD 2520
 80 DSAVILGPPG PRGAKGDMGE RGPRLDGD GPRGDNPDG DKSGKEPGD KGSAGLPLGR 2580
 GLLGPGQPG AAGIPGDPGS PGKDGVPGR GEKGDVFGM PRGLKGERGV KGACGLDGEK 2640
 GDXGAGPGE RPLAGHKGE MEGPVPQGS GAPGKEGLIG PKGDRGFDQ GPKGDKGEK 2700
 GERGTPGIGG PPGPSGNDGS AGPPGPPGSP GPRGPEGLQ KGERGPPGE RVVGPAGV 2760
 APGERGEQGR PPGAPRGEK GEALTEDDI RGFVRQEMSG HCACQGGFIA SGRSLPSPYA 2820
 85 ADTAGSQLHA VPVLRVSHAS EERVPFEDD EYSEYSEY EYQDPEAPW DSDDPCLSLP 2880
 DEGSCTAYTL RMYHRAVTS TEACHPFVYG GCGGNANRFG TREACERRCP PRVVQSGGTG 2940
 TAQD

	1	11	21	31	41	51	
25	TAGTCCGGAGC	GAGGTGGCGA	GTCGCTGAGC	CCGCCCGCGC	CCCGAGAGCG	GCTGCGACCG	600
	CCCGCGCCGG	GAGGAGGAGG	GCGAGGCGCG	CCCGAGCCCG	CGCCCGCCGC	GGCCACCGCG	1200
	CCCGCGCCAC	CAGCCGCCAC	GAGCTCGCGG	GCCAGCCGGG	CAGCCTCGCG	GGCGCCCGCG	1800
	CGGGGCGGGG	GGCGCGGGCC	ACAGGCCCTC	GCTCCGCGCG	TGCTTTGTCAG	ACCGCGGGCG	2400
30	CCGATGTCCG	CCGCGCCCCG	TAGAGATGAT	TCTGGGTGCG	CGCGGAGGAG	CCCGCGACGC	3000
	GGCGTCGCGC	GAGCGCGGGG	CAGGAGCGCT	GGGAGCGCCG	GGCGCGCCGC	CGCGCGCCGC	3600
	GGCGGCGCCC	GACGCGCGCC	CGCGCGCCCC	GGGCCCCGCA	CACACATGAG	ATTCTTCCAG	4200
	CTCACITTTA	AGTGTCTCGT	GSACTTTCGT	TGACTGTGCG	GGCGCGCCGC	CGACCCCGCG	4800
35	CGTCCGCCCC	CGCGCCCGTC	CCCGCGCGCG	CGCGCGCCCC	GGCCCCCGGC	CGCGCGCGCG	5400
	CCTCGGGGCG	TCTCCCCGTC	CGCGCGGTGC	CCCCCGCCCT	ACCGCGCGCC	CCGCTGAGCG	6000
	CGCGCGACCG	CGCGCGCGCC	GTCGCGCGCG	CGCGGGCAT	GGCGAAGAGT	ACCGCGGAGA	6600
	ACCGCATCTA	TAGCTGTCTC	GGCGACCGAG	AGAAGGGCCC	CCTCATCGCG	CCCGGGCGCG	7200
	ACGGGCGGCG	GGCGAAGGGG	GACGGCGGCG	TGGGGCTGGG	GACACCCGCG	GGCGCGCTGG	7800
40	CCGTGCGCCG	CGCGGAGACG	TGGACCGCGC	AGATGTGACT	CATCATGTAG	TGCTTGGGCT	8400
	TGCGCGTGGG	CTTGGGCAAC	GTTGTGGCGT	CCCTCTACT	GCTGTACIAG	AAAGCGGGAG	9000
	GTGTGTTTCT	TATTCCCTAC	GTCCTGATCG	CCCTGTTTGG	AGGAATCCCC	ATTTTCTTCT	9600
	TAGAGATCTC	GCTGGGCCAG	TTCATGAAGG	CGCGGACAT	CAATGTCTCG	AAACATCTGT	10200
	CCGCTTTCAA	AGGCTCTGGC	TACGGCTCTA	TGGTGTGCT	CTTCTACTCG	AAACCTTACT	10800
45	ACATCATGCT	GCTGGCGCTG	GGCTTCTTAT	ACCTGTGTCA	GCTCTTACCC	AGCCACCTGC	11400
	CTTGCGGACG	ATGTGGCCAC	CTGCTGGAAC	TCCCGACTCG	GTGCGAGATC	TTCCGCGAGT	12000
	AAGACTGTGC	CAATGCCAGC	CTGGCCCAAC	TCACCTGTGA	CCAGCTTGCT	GACCGCGGCT	12600
	CCCGCTGTCT	CGAGTCTCTG	GAGACAGACT	TCTTGAGGCT	GCTTGGGGGA	CTGGAGGTGC	13200
	CAGGGGCCCT	CAACTGGGAG	TGACACCTTT	GTCGTCTGCG	GTGCTGGGTA	GTGGTCTACT	13800
50	CTGTGCTGTG	GAAGGGGGTC	AAATCTGCGG	GAAAGATCGT	GTACTTCACT	GCTCATATCC	14400
	CTCATGTGCT	CTGTGTGCTG	CTTGTGTGTC	TGGAGTGTCG	GTCGCTGGC	GGCCGTGATG	15000
	GCATCATTTA	CTATCTCAAG	CTCGTACGTG	CAAAGCTGGG	GTCGCCCTCA	GTGTGGATAG	15600
	ATGCGGGGAC	CCAGATTTTC	TTTTCTTACG	CCATTGCCCT	GGCGGCCCTC	ACGACCCCTG	16200
	GCAGCTACAA	CGCTTCAAC	AACACATGCT	ACAAGGACGC	CATCATCTCG	GCTCTCATCA	16800
55	ACAAGTGGAC	CAGCTTCTTT	GCTGGGTCTG	TGGTCTTTCT	CATCTCTGGC	TTTATGTGCT	17400
	CAGACGAGCG	GTCGCACATC	TCCAAAGTGG	CAGAGTCAGG	GGCGGGCTCG	CGCTTCATCG	18000
	CCTACCGCGG	GGCTGTCAAG	CTGATGCCAG	TGGCCCCACT	CTGGGCTGCC	CTTGTCTTCT	18600
	GCTCTCTGTT	GCTGCTTGGT	CTCGACAGCG	TGTTGTAGAG	GTGCGAGGCG	TTTATCATCG	19200
	TCGCTCTGTA	CGCTCTCGCG	GGCTCTTACT	ACTTTCGTTT	CCAAAGGGAG	ATCTCTGTGG	19800
60	CCCTCTGTGT	TGCCCTCTGC	TGTTGTCATG	ATCTCTCATC	GATGTACTAG	GGCGGGAGTG	20400
	ACGCTTTCCA	GCTGTTTGAC	TACTACTCGG	CCAGCGGCAC	CCGCTGCTC	TGGCAGAGCT	21000
	TTTGGGAGTG	GCTGGTGGTG	GCTGGGTTGT	ACCGAGCTGA	CGGCTTCAAT	GACGACATTG	21600
	CTGTATGATG	GGGTACCGCA	CCCTGGCCCT	GGATGAAATG	GTGCTGGTCT	TGTCATCTAC	22200
	CGCTGGTCTG	CATGGGCATC	TTTCACTTCA	ACGTTGTGTA	CTACGAGCGC	CTGGTCTTCA	22800
65	ACAACCACTA	CGTGTGCTG	TAGGTTGGGT	AGGCCATGGG	CTGGGCCCTC	GGCCCTTCCT	23400
	CATAGCTGCG	GTCGGCGGCT	CACCTCTCTG	GTCGCCCTCT	CAGGGCCCAAG	GGACACATGG	24000
	CTGAGCGCTG	CAGCACCTCT	AGCCACGCCA	TCTGGGGGCT	CACACCACTG	GAGTACCGAG	24600
	CTCAGGACCG	AGATGTCAAG	GGCCCTGACCA	CCCTGACCCC	AGTGTCCGAG	AGCAGCAAGG	25200
	TGCTGTGTGT	GGAGAGTGTG	ATGTGACAC	TCAGCTTACA	TCACACAGCT	ACCTCTGGTA	25800
70	GGCATAGACG	CGGCTGCTTC	AGGCCCATCT	CAGCCCTCCA	GGGGGCCCTG	CTTTCCTTGA	26400
	CACTTTTCGT	GTCTCGGTGG	GGGAGGAGGG	GAGAAAGCAC	CATGAGTGCT	CATTAAACAA	27000
	ACTTTTTCCT	TTTTTATAAA	AAAGCCCAAA	ATATCAACAC	CCACCAAAAG	TAGATGCTCT	27600
	TCCCCCTACA	GGCCTAGCGC	AGCTGGTCTT	AGSGCCCCCG	TAGTGGCCCC	CCCCCAACCA	28200
	CAGTGTGCTA	CTCCTCTCTG	CCCTGCCACG	CCACCCCCCT	GGCCACCTCT	CAGGCTCTTG	28800
75	CTCTCGAGCA	CACCCGTGGG	TGACCCCTCA	CCCCAGAGCA	AGCAGTGGCA	CTTGTGGGAA	29400
	TGTGAGGAAG	GGAGGAGGAG	AGAGGACGGG	GGGAGGAGAG	AGAGGAGAAG	GGAGGCAGGG	

CTTCTGTGTA GCAGCTTTAA CCCAGTTTG TCTGTCACT CCAGTCCCGA GACGGCTGAG 3780
 TGACCCCAAG AAAGGCTTCC CCGACACCCA GACAGAGGCT GCAGGGCTGG GCTGGGTGA 3840
 GGGTGGCGGG CCTGGCGGGA CATTCTACTG TGCTAAAAAG CCAGTGCAGA CATAGCAATA 3900
 AAAACATGTC ATTTTCC

Seq ID NO: 67 Protein sequence:
 Protein Accession #: NP_005620.1

1 11 21 31 41 51
 10 | | | | |
 MAKKSARENGI YSVSGDEKKK PLIAPGPDGA PAKGDGPVGL GTPGGRRLAVP PRETWTRQMD 60
 PIMSCVGFPAV GLGNVWRFPY LCYKNGGGVF LIPYVLIALV GGIPIFLEI SLGQFMKAGS 120
 INVWNICPLF KGLGYASPMI VFYCNYYIIM VLANGFYVLV KSFTTTLFPA TCGHTWNTPD 180
 15 CVELFRHEDC ANASLANLTC DQLADRRSPV IEFWENKVLV LSGGLEVPGA LNWEVTLCLL 240
 ACWVLVYFCV WKGVKSTGKI VYPTATFPYV VLVVLLVRGV LLPGALDGI IYLLKPDMSKL 300
 GSPQVWIDAG TQIFFSYAIG LGALTALGSY NRFNNNNCYK AIIILALINS TSPFAGPVFV 360
 SILGFMMAEG GVHISKVAES GPGLAFIAYP RAVTLMPVAP LWALFFPML LLLGLDSQPV 420
 GVEGFITGLL DLLPASYYFR PQREISVALC CALCFVIDLS MVTGGMYVF QLPDYYSASG 480
 20 TLLWQAFWE CVVVAWVYGA DRFMDIACM IGYRCPFWMK WMSFPTPLV CMGIFIFNVV 540
 YVEPLVYNT YVYPWGEAM GWAFALSSML CVPLHLLGCL LRKGTMAER WQHLTQPIWG 600
 LHLLEYRAQD ADVRGLTILT PVSESSKVYV VESVM

Seq ID NO: 68 DNA sequence.
 Nucleic Acid Accession #: NM_021953.1
 Coding sequence: 178-2469

1 11 21 31 41 51
 30 | | | | |
 GGCACGAGGG GGACCCGCGC GGTCCGGCGC GAGCCCCCGT CCGGGGCCCT GGTCCGGCCC 60
 CCAGGTTGGA GGAGCCCGGA GCCCGCCTTC GGAGTACGG CCTAACGGCG GCGGCGACTG 120
 CAGTCTGAGG GGTCCACACT TGTGATTCTC AATGGAGAGT GAAACGCGAG ATTCATAATG 180
 AAAGCTAGCC CCGCTGGGCC ACTGATTCTC AAAAGACGGA GGCTGCCCTC TCCTGTTCAA 240
 AATGCCCCAA GTGAACATC AGAGAGGAGAA CCTAAGAGAT CCGCTGCCCA ACAGGAGTCT 300
 35 AATCAAGCAG AGGCCCTCAA GGAAGTGGCG GAGTCCAAC CTTCGAAGTT TCCAGCTGGG 360
 ATCAAGATTA TTAACCAACC CACCATGCCG AACACGCAAG TAGTGGCCAT CCCCACAAT 420
 GCTAATATTC ACAGCATCAT CACAGCACTG ACTGCCAAGG GAAAAGAGAG TGGCAGTAGT 480
 GGGCCCAACA AATTCTCTCT CATCAGCTGT GGGGAGGCC CAACTCAGCC TCCAGGACTC 540
 CGGCCTCAAA CCCAAACGAG CTATGATGCC AAAAGGACAG AAGTGACCTT GGAGACCTTG 600
 40 GGACCAAAAC CTGCAGCTAG GATGTGAAT CTTCCTAGAC CACCTGGAGC CTTTGGCAG 660
 CAGAACACGG AGACCTGTGC AGATGGTAGG GCAGCAGGCT GCACTATCAA CAATAGCCTA 720
 TCCAACATTC AGTGGCTTCG AAAGATGAGT TCTGATGGAC TGGGCTCCCG CAGCATCAAG 780
 CAGAGATGAG AGGAAAGGGA GAATTGTCTC CTGGAGCAGC GACAGGTTAA GGTGAGGAG 840
 CTTTCGAGAC CATCAGCTGC CTGGCAGAAC TCTGTGTCTG AGCGGCCACC CTACTCTTAC 900
 45 ATGGCCATGA TACAATTCGC CATCAACAGC ACTGAGAGGA AGCGCATGAC TTTGAAAGAC 960
 ATCTATAGCT GGATTGAGGA CCACCTTCCC TACTTTAAGC ACATTGCCAA GCCAGGCTGG 1020
 AAGAACTCAG TCCGCCACAA CCTTTCCTCG CACGACATGT TTGTCCGGGA GAGCTCTGCG 1080
 AATGGCAATG TCTCTCTCTG GACCATTCAC CCGAGTGCCA ACCGCTACTT GACATTTGAC 1140
 CAGGTGTTTA AGCCACTGCA CCGAGGTCTT CCACAATTGC CGAGCAGACT GGAATCAGAC 1200
 50 CAGAAACGAC CGAATCCAGA GCTCCGCGCG AACATGACCA TCAAAACCGA ACTCCCCCTG 1260
 GGGCAGCGCG GGAAGATGAA GCCACTGCTA CCACGGGTCA GCTCATACCT GGTACCTATC 1320
 CAGTTCCCGG TGAACCACTC ACTGGTGTTC CAGCCCTCGG TGAAGTGCC ATTGCCCTAG 1380
 CGGCTTCCCT TCATGAGCTC AGAGCTTGCC GCCCATAGCA AGGAGTCCG CATTCGCCCT 1440
 AAGGTGCTGC TAGCTGAGGA GGGGATAGCT CCTCTTCTT CTGAGGACC AGGAAAGAG 1500
 55 GAGAAACTCT TGTGAGAGA AGGTTTCTCT CCTTGTCTC CAGTTCAGAC TATCAAGGAG 1560
 GAGAAATTC AGCCTGGGGA GGAATGCCA CACTTAGCGA GACCCATCAA AGTGGAGAGC 1620
 CCTCCCTTGG AAGAGTGCCG CTCCCGGCC CCATCTTTCA AAGAGGAATC ATCTCACTCC 1680
 TGGGAGGATT CGTCCCAATC TCCCAACCCA AGACCCAAGA AGTCCCTACG TGGGCTTAGG 1740
 TCCCAACCCC GGTGTGTCTC GGAATGCTT GTGATTCAAC ACAGGAGAGG GAGGGAGAGG 1800
 60 AGCCCGTCTC GGAGGAAACA GCATCTACTG CCTCCCTGTG TGGATGAGCC GGAGCTGCTC 1860
 TTCTCAGAGG GGCCAGTAC TTCCCGCTGG GCCGAGAGC TCCCGTTCCT AGCAGACTCC 1920
 TCTGACCTCG TCTCCAGCTC CAGCTACTCC CAGGAAGTGG GAGGACCTTT TAAGACACCC 1980
 ATTAAGGAAA CGCTGCCACT CTCTCCACC CCGAGCAAT CTGTCTCCCT CAGAACCCTT 2040
 GAATCTGGA GGTCTACGCC CCCAGCCAAA GTAGGGGGAG TGGATTTCAG CCCAGTACAA 2100
 65 ACCTCCAGG GTGCTCTGA CCGCTTGCTT GACCCCTTGG GGTGATGGA TCTCAGCACC 2160
 ACTCCCTTGC AAGTGTCTCC CCCCCTTGAA TCACCGCAAA GGCTCCTCAG TTCAGAACCC 2220
 TTAGACCTCA TCTCCGTCCC CTTTGGCAAC TCTTCTCCCT CAGATATAGA CGTCCCAAG 2280
 CCAGGCTCCC CGAGGCCACA GGTTCCTGGC CTGTCAGCCA ATCGTTCTCT GACAGAAGGC 2340
 CTGGTCTCGG ACACAATGAA TGACAGCCTC AGCAAGATCC TGCTGGACAT CAGCTTCTCT 2400
 70 GGCTTGGAGC AGGACCCACT GGGCCCTGAC AACATCAACT GGTCCCAGTT TATTCCTGAG 2460
 CTACAGTAGA GCCCTGCCCT TGCCCTGTGT CTCAAGCTGT CCACCATCCC GGGCACTCCA 2520
 AGGCTCAGTG CACCCCAAGC CTCTGAGTGA GGACAGCAGG CAGGGACTGT TCTGCTCCTC 2580
 ATAGCTCCCT GCTGCTGAT TATGCAAAAG TAGCAGTCAC ACCCTAGCCA CTGCTGGGAC 2640
 CTTGTGTTCC CCAAGAGTAT CTGATTCTCT TGCTGTCCCT GCCAGGAGCT GAAGGGTGGG 2700
 75 AACAAACAAG GCAATGGTGA AAAGAGATTA GGAACCCCCC AGCTGTGTTT CATCTCTGTC 2760
 CCAGCAGTCT CTACCTTTC CTGATCTTTG CAGGGTGGTC CGTGTAATA GTATAAATTC 2820
 TCCAAATTTA CTTCTAATTA TAAATGTAA CTTATTCTCT TAGATCATT TCCAGAGACT 2880
 GCCAGAGGTT GGGTAGGATG ACCTGGGGTT TCAATTGACT TCTGTCTCTT GCTTTTAGTT 2940
 TTGATAGAA GGAAGACCTG CAGTGACAGG TTTCTTCCAG GCTGAGGTAC CTGATCTTG 3000
 80 GGTCTTTCAC TGCAGGGACC CAGACAAGTG GATCTGCTTG CCAGAGTCTT TTTTGGCCCT 3060
 CCCTGCCACC TCCCGGTGTT TCCAAGTCAG CTTCCTGCA AGAAGAAATC CTGGTTAAAA 3120
 AAGTCTTTTG TATTGGGTCA GGAGTTGAAT TTGGGGTGGG AGGATGGATG CAACCTGAAGC 3180
 AGAGTGTGGG TGCCAGATG TGCGCTAAT GATGTTCTC TGATAATGTC CCAATCATA 3240
 CCAGGGAGAC TGCATTGAC GAGAATCAG GTGGAGGCTT GAGAAGGCCG AAAGGGCCCC 3300
 85 TGACCTGCTT GGCTTCTCTA GCTTGCCCTT CAGCTTTGCA AAGAGCCACC CTAGGCCCCA 3360
 GCTGACCGCA TGGGTGTGAG CCAGCTTGAG AACACTAAT ACTCAATAAA AGCGAAGGTG 3420
 GACCAAAAAA AAAAAA AAAA

Seq ID NO: 69 Protein sequence:
Protein Accession #: NP_068772.1

5 1 11 21 31 41 51
MKASPRRLPI LKRRRLPLFV QNAPSETSEE EPKRSPAQQE SNQAEASKEV AESNSCKFPA 60
GIKIIINHPTM PNTQVVAIFN NANIHSIITA LTAKGKESGS SGPKNKILIS CGGAPTQPPG 120
LRPQTQTSYD AKRTEVTLET LGPKPAARDV NLPRPPGALC DQKRETCADG EAAGCTINNS 180
10 LSNITQNLKRM SSDGLGSRSI QEMEEKENC HLEQRQVKVE EPSRPSASWQ NSVSEPPYS 240
YMAMIQFAIN STERKRM TLK DIYTWIEDHF PYFKHIAKPG WKNSIRHNLS LHDMFVRETS 300
ANGKVSFWTI HPSANRYLTL DQVFKPLDPG SPQLPEHLES QKRRPNPELR RNMTIKTELP 360
LGARRKMKPL LPRVSSYLVP IQFFVNVQSLV LQPSVKVPLP LAASLMSEL ARHSKRVRIA 420
PKVLLAEBGI APLSSAGPGK EEKLLFGEGF SPLLPVQTIK EEEIQGEEM PHLARPIKVE 480
15 SPLEEWSP APFSKKESSH SWEDSSQSPT PRPKSYSGL RSPTRCVSEM LVIQHRERRE 540
RSRRRKQHL LPPCVDPEPL LFSEGPSTSR WAAELPPAD SSDPASQLSY SQEVGGPFRT 600
PIKETLPISS TPKSVLPRT PESWRLTPPA KVGGLDFSPV QTSQASDPL PDLGLMDLS 660
TTPLOAPPL ESPORLLSE PLDLISVFPF NSSPSDIDVP KGPSPEPQVS GLAANRSLTE 720
20 GLVLDTMNDS LSKILLDISF PGLDEDPLGP DNINWSQFIP ELQ

Seq ID NO: 70 DNA sequence
Nucleic Acid Accession #: BC006529.1
Coding sequence: 178-2424

25 1 11 21 31 41 51
GGCACGAGGG GGACCCGGCC GGTCCGGCGC GAGCCCCCGT CCGGGGCCCT GGCTCGGCC 60
CCAGGTGGGA GGAGCCCGGA GCCCGCCTTC GGAGCTACGG CCTAACGGCG CGCGCGACTG 120
30 CAGTCTGGAG GGTCCACACT GTGATTCTC AATGGAGAGT GAAAACGACG ATTCATAATG 180
AAAACCTAGCC CCGCTCGGCC ACTGATTCTC AAAAGACGGA GGCTGCCCTT TCCTGTTCAG 240
AATGCCCAAA GTGAAACATC AGAGGAGGAA CCTAAGAGAT CCCCTGCCCA ACAGGAGTCT 300
AATCAAGCAG AGCCCTCCAA GGAAGTGGCA GAGTCCAACT CTGTCAAGTT TCCAGCTGGG 360
ATCAAGATTA TTAACCAACC CACCATGCCC AACACGCAAG TAGTGGCCAT CCCCACAAT 420
35 GCTAATATTC ACAGCATCAT CACAGCACTG ACTGCCAAGG GAAAAGAGAG TGGCAGTAGT 480
GGGCCCAACA AATTCTATCT CATCAGCTGT GGGGGAGCCC CAACTCAGCC TCCAGGACTC 540
CGGCCTCAAA CCCAACACAG CTATGATGCC AAAAGGACAG AAGTGACCTT GGAGACCTTG 600
GGACCAAAAC CTGCAGTAGG GGATGTGAAT CTCTCTAGAC CACCTGGAGC CCTTTGCGAG 660
CAGAAACGGG AGACCTGTGC AGATGGTGAG GCAGCAGGCT GCACTATCAA CAATAGCCTA 720
40 TCCAACATCC AGTGGCTTCG AAGATGAGT TCTGATGGAC TGGGCTCCCG CAGCATCAAG 780
CAAGAGATGG AGGAAAAGGA GAATTGTAC CTTGAGCAGC GACAGGTAA GGTGTAGGAG 840
CCTTCGAGAC CATCAGCGTC CTGGCAGAAC TCTGTGCTGT AGCGGCCACC TACTCTTAC 900
ATGGCAGTGA TACAATTGCG CATCAACAGC ACTGAGAGGA AGCGCATGAC TTTGAAAGAC 960
45 ATCTATAGT GGAATGAGGA CCACCTTCCC TACTTTAAGC ACATTGCCAA GCCAGGCTGG 1020
AAGAACTCCA TCCGCCACAA CCTTCCCTG CACGACATGT TTGTCCGGGA GACGTCTGCC 1080
AATGGCAAGG TCTCTTCTG GACCATTCAC CCCAGTGCCA ACCGCTACTT GACATTGGAC 1140
CAGGTGTTTA AGCAGCAGAA ACAGACGAAAT CCAGAGCTCC GCCGGAAACT GACCATCAA 1200
ACCGAACTCC CCTCGGGCGC ACGGCGGAAG ATGAAGCCAC TGCTACCAAG GGTCACTCA 1260
50 TACCTGGTAT CTATCCAGTT CCGGTGAAC CAGTCACTGG TGTTCAGCC CTGGGTGAAG 1320
GTGCCATGTC CCTCGGCGGC TCCCTCATG AGCTCAGAGC TTGCCCGCCA TAGCAAGCGA 1380
GTCCGCTATT CCCCCAAGGT GCTGCTAGCT GAGGAGGGGA TAGCTCCTCT TTCTCTGCA 1440
GGACCAAGGA AAGAGGAGAA ACTCCTGTCT GGAGAGGGGT TTCTCCTTT GCTTCCAGTT 1500
CAGACTATCA AGGAGGAAGA AATCCAGCCT GGGGAGGAAA TGCCACACTT AGOGAGACCC 1560
55 ATCAAGATGG AGAGCCCTCC CTGGAAGAG TGGCCCTCCC CGGCCCATC TTTCAGAGAG 1620
GAATCATCTC ACTCTGGGA GGATTCGTCC CAATCTCCA CCCCAGACC CAAGAAGTCC 1680
TACAGTGGCC TTAGGTCCCC AACCCGCTGT GTCTCGGAAA TGCTTGTGAT TCAACACAGG 1740
GAGAGGAGGG AGAGGAGCGT GTCTCGGAGG AAACAGCATC TACTGCTCC CTGTGTGGAT 1800
GAGCGGAGAG TGTCTTCTC AGAGGGGCCC AGTACTTCCC GCTGGGCGCG AGAGCTCCCG 1860
60 TTCCAGCAG ACTCTCTGA CCTGCTCTC CAGCTCAGCT ACTCCAGGA AGTGGGAGGA 1920
CCTTTAAGA CACCATTAA GGAAACGCTG CCCATCTCCT CCACCCGAG CAAATCTGTC 1980
CTCCCCAGAA CCCCTGAATC CTGAGGCTC ACGCCCCAG CCAAGTAGG GGGACTGGAT 2040
TTCAGCCAG TACAACCCCT CCAGGCTGCC TCTGACCCCT TGCCCTGACC CTGCGGCTG 2100
ATGATCTCA GCACCACTCC CTTCGAAAGT GCTCCCCCCC TTGAATCACC GCAAGGCTC 2160
CTCAGTTTCA AACCCCTAGA CCTCATCTCT GTCCCTTTG GCAACTCTTC TCCTCTCAGT 2220
65 ATAGACGTC CCAAGCCAGG CTCCCGGAG CCACAGGTTT CTGGCCTTGC AGCCAATCGT 2280
TCTCTGACAG AAGGCTGTGT CCTGGACACA ATGAATGACA GCCTCAGCAA GATCCTGCTG 2340
GACATCAGCT TTCTTGCCCT GGACGAGGAC CCACTGGGCC CTGACACAT CAACTGGTCC 2400
CAGTTTATTC CTGAGCTACA GTAGAGCCCT GCCCTTGCCC CTGTGCTCAA GCTGTCCACC 2460
70 ATCCCGGGCA CTCCAGGCT CAGTGACCCC CAAGCCTCTG AGTGAAGACA GCAGGCGAGG 2520
ACTGTTCTGC TCCTCATAGT TCCTTGCTGC CTGATTATGC AAAAGTAGCA GTCACACCT 2580
AGCCACTGCT GGGACCTGT GTTCCCAAG AGTATCTGAT TCCTCTGCTG TCCTTGCCAG 2640
GAGCTGAAG GTGGGAACAA CAAAGGCAAT GGTGAAAAGA GATTAGAAC CCCCAGCCT 2700
75 GTTTCCTTC TCTGCCAGC AGTCTCTTAC CTTCCTGAT CTTTGCAAGG TGGTCCGTG 2760
AAATAGTATA AATTCTCCAA ATTATCTCTT AATTATAAT GTAAGCTTAT TTCTTAGAT 2820
CATTATCCAG AGACTGCCAG AAGTGGGTA GGATGACCTG GGGTTTCAAT TGACTTCTG 2880
TCCTTGCTTT TAGTTTGTAT AGAAGGGAAG ACCTGCAGTG CACGGTTTCT TCCAGGCTGA 2940
GGTACCTGGA TCTTGGGTTT TTCACTGCAG GGACCCAGAC AAGTGGATCT GCTTGCCAGA 3000
GTCCCTTTTG CCCTCTCCTG CCACCTCCCT GTGTTTCAA GTCACTCTTC CTCGAAGAAG 3060
80 AAATCCTGTT TAAAAAGTC TTTGTATTG GTTCAGGAGT TGAATTTGGG GTGGGAGGAT 3120
GGATGCACT GAAGCAGAGT GTGGGTGCCC AGATGTGCGC TATTAGATGT TTCTCTGATA 3180
ATGTCCCAAA TCATACCAGG GAGACTGGCA TTGACGAGAA CTCAGTGGGA GGCTTGAGAA 3240
GGCCGAAGG GCCCTGACC TGCCGTGCTT CTTAGCTTG CCCCTCAGCT TTGCAAGAG 3300
CCACCTAGG CCCCAGCTGA CCGCATGGGT GTGAGCCAGC TTGAGAACAC TAACACTCA 3360
85 ATAAAGGCGA AGGTGAAAAA AAAAAAAA AAAAAA

Seq ID NO: 71 Protein sequence:
Protein Accession #: AAH06529.1

1 11 21 31 41 51
5 MKTSPRRPLI LKRRRLPLFV QNAPSETSEE EPKRSAPAQZE SNQAEASKEV AESNSCKFFPA 60
GIKIIINHPTM PNTQVVAIFN NANIHSLITA LTAKGKESGS SGPNKPILIS CGGAPTQPPG 120
LRPOTQTSYD AKRTEVTLST LGPKPAARDV NLPRPPGALC EQKRETCADG BAAGCTINNS 180
LSNIQWLKRM SSDGLGSRSI QDEMEEKENC HLEQROVKVE EPSRPSASWQ NSVSEPPYS 240
YMAHQFAIN STERKMTLK DIYTWIEDHE PYFKHIAKPG WKNSIRHNL LHMDFVRETS 300
10 ANGKVSFWTI HPSANRYLTL DQVFKQKRE NPFLRRNMTI KTELPLGARR EMKPLLPVRS 360
SYLVPIQFPV NQSLVLQPSV KVPLFLAASL MSSELARESK RVRIAPKVL AEESGIAPLSS 420
AGPGKEEKL FEGGFSPLLP VQTIKEEEIQ PGEEMPHLAR PIKVESPPLE EWSPAPSPFK 480
EESHSWEDS SQSPTRPKK SYSGLRSPTR CVSEMLVIQH RERRERSRSR RKQHLPLPCV 540
DEPELLFSEB PSTSRWAAEL PFPADSSDPA SQLSYSQEVG GPFRTPIKET LPISSTPSKS 600
15 VLPRTPEWR LTPPAKVGLG DFPSPVTPQG ASDPLPDLG LMDLSTPLQ SAPPLESPQR 660
LLSSEPLDLI SVPPGNSSPS DIDVPKPGSP EPQVSGLAAN RSLTEGLVLD TMNDSLSKIL 720
LDISFPGLE DPLGPDNIW SQFIPELQ

20 Seq ID NO: 72 DNA sequence
Nucleic Acid Accession #: U74612.1
Coding sequence: 178-2583

1 11 21 31 41 51
25 GGCACGAGGG GGACCGGGCC GGTCCGGCGC GAGCCCCCGT CCGGGGCCCT GGCTCGGCC 60
CCAGGTGGA GGAGCCCGGA GCCCGCCTTC GGAGCTACGG CCTAACGGCG CGCGGACGTG 120
CAGTCTGGAG GGTCCACACT TGTGATTCTC AATGAGAGT GAAACGCGAG ATTCATAATG 180
AAAACCTAGCC CCGTGGGGCC ACTGATTCTC AAAAGACGGA GGCTGCCCCT TCCTGTTCAA 240
AATGCCCAA GTGAAACATC AGAGGAGGAA CCTAAGAGAT CCGCTGCCCA ACAGGAGTCT 300
30 AATCAAGCAG AGGCTCCCAA GGAAGTGCCA GAGTCCAACT CTGCAAGTT TCCAGCTGGG 360
ATCAGATTA TTAACACACC CACCATGCCC AACACGCAAG TAGTGGCCAT CCCCACAAT 420
GCTAATATTC ACAGCATCAT CACAGCACTG ACTGCCAAGG GAAAGAGAG TGGCAGTAGT 480
GGGCCCAACA AATTCATCCT CATCAGCTGT GGGGGAGCCC CAACTCAGCC TCCAGGACTC 540
CGGCCTCAA CCCAACACG CTATGATGCC AAAAGGACAG AAGTGACCTT GGAGACCTTG 600
35 GGACCAAAAC ATGCAGCTAG GGAATGGAAT CTTCTAGAC CACCTGGAGC CTTTGGCAG 660
CAGAAACGGG AGACCTGTGC AGATGTTGAG CGACGAGGCT GCACTATCAA CAATAGCCTA 720
TCCACATCC AGTGGCTTCG AAAGATGAGT TCTGATGGAC TGGGCTCCCG CAGCATCAAG 780
CAAGAGATGG AGGAAAGGGA GAATTGTCAC CTGGAGCAGC GACAGGTTAA GGTGAGGAG 840
CCTTCGAGAC CATCAGCTGC CTGGCAGAAC TCTGTGCTG AGCGGCCACC TACTCTTAC 900
40 ATGGCCATGA TACAATTCCG CATCAACAGC ACTGAGAGGA AGCGCATGAC TTTGAAAGAC 960
ATCTATAGT GGAATGAGGA CCCTTTCCCT TACTTTAAGC ACATTGCCAA GCCAGCTGG 1020
AAGAACTCCA TCCGCCACAA CCTTTCCCTG CACGACATGT TTGTCGGGA GACGTCTGCC 1080
AATGGCAAGG TCTCCTCTG GACCATTCAC CCCAGTGCCA ACCGCTACTT GACATGGAGC 1140
CAGGTGTTA AGCCACTGGA CCCAGGCTCT CCACAAATGC CGGAGCACTT GGAATCAGC 1200
45 CAGAAACGAG CGAATCCAGA GCTCCGCGGG AACATGACCA TCAAAACGGA CCTTCCCTG 1260
GGCGCACGGC GGAAGATGAA GCCACTGCTA CCACGGGTCA GCTCATACCT GGTACCTATC 1320
CAGTTCCCGG TGAACAGCTG ACTGGTGTG CAGCCCTCGG TGAAGGTGCC ATTGCCCTG 1380
GCGGCTTCCC TCAATAGCTC AGAGCTTGCC CGCCATAGCA AGCGAGTCCG CATTTGCCCTC 1440
AAGGTTTTTG GGGAAACAGT GGTGTTTGGT TACATAGATA AGTTCTTTAG TGGGATCTG 1500
50 CGAGATTTTG GTACACCCAT CACCAGCTTG TTTAATTTA TCTTCTTTG TTTATCAGT 1560
CTGCTAGCTG AGGAGGGGAT AGCTCTCTCT TCTTCTGCG GACCAAGGAA AGAGGAGAAA 1620
CTCTCTTTG GAGAAGGTTT TTCTCCTTTG CTTCCAGTTC AGACTATCAA GGAGGAAGAA 1680
ATCCAGCTG GGGAGGAAAT GCCACACTTA GCGAGACCCA TCAAGTGGA GAGCCTCCC 1740
TTGGAAGAGT GGGCCTCCCC GGGCCCATCT TCAAGAGAG AATCATCTCA CTCTGGGAG 1800
55 GATTCTGCC CCAATCCAC CCAAGAGCCC AAGAAGTCTC ACAGTGGGCT TAGGTCCCCA 1860
ACCGGTGTG TCTCGGAAT GCTTGTGATT CAACACAGGG AGAGGAGGGA GAGGAGCCGG 1920
TCTCGGAGGA AACAGCATCT ACTGCCCTCC TGTGTGGATG AGCCGAGGCT GCTCTTCTA 1980
GAGGGGCCCA GTACTTCCCG CTGGGCGGCA GAGCTCCCGT TCCAGCAGA CTCTCTGAC 2040
CCTGCTCCC AGCTCAGTCA CTCCAGGAA GTGGGAGGAC CTTTAAGAC ACCCATTAAG 2100
60 GAAACGCTGC CCAATCTCTC CACCCGAGC AATCTGTGCC TCCCGAGAAC CCTGATTC 2160
TGGAGGCTCA CGCCCCAGC CAAAGTAGGG GAGCTGGATT TCAGCCAGAT ACAACCTCC 2220
CAGGGTGCTC CTGACCCCTT GCGTACCCCC CTGGGCTGA TGGATCTCAG CACCACTCC 2280
TTGCAAGTGT CTCCTCCCCC TGAATCACCG CAAAGGCTCC TCAGTTCAGA ACCCTTAGAC 2340
CTCATCTCCG TCCCTTTTGG CAATCTCTCT CCTCAGATA TAGAGTCCC CAGCCAGGC 2400
65 TCCCGGAGC CACAGGTTTC TGGCCTTGCA GCCAATCGTT CTCTGACAGA AGCCCTGGTC 2460
CTGACACAA TGAATGACAG CCTCAGCAAG ATCCTGCTGG ACATCAGCTT TCCTGGCCTG 2520
GACGAGGACC CACTGGGCCC TGACAACATC AACTGTGCCC AGTTTATTC TGAGCTACAG 2580
TAGAGCCCTG CCGTGTCCCG TGTGCTCAAG CTGTCCACCA TCCCGGGCAC TCCAAGGCTC 2640
AGTGACCCCC AAGCCTCTGA GTGAGGACAG CAGGACGGGA CTGTTCTGCT CCTCATAGCT 2700
70 CCGTCTGCC TGATTATGCA AAGTAGCAG TCACACCTTA GCCACTGCTG GGACCTGTG 2760
TTCCCAAGA GTATTCTGAT CCTCTGCTGT CCGTCCAGG AGCTGAAGGG TGGGAACAAC 2820
AAAGGCAATG GTGAAAGAG ATTAGGAACC CCCCAGCCTG TTTCCATCT CTGCCACGA 2880
GTCTCTTACC TTCCCTGATC TTTGAGGGT GGTCCGTGTA AATAGTATA ATTCTCCAA 2940
TTATCTCTA ATTATAAATG TAAGCTTATT TCCTTAGATC ATTATCCAGA GACTGCCAGA 3000
75 AGGTGGGTAG GATGACCTGG GGTTCATATT GACTTCTGTT CTTGCTTTT AGTTTGATA 3060
GAAGGGAAGA CTTGAGTGC ACGGTTTCTT CCAGGCTGAG GTACCTGGAT CTGCGGTCT 3120
TCACTGCAGG GACCCAGACA AGTGGATCTG CTTGCCAGAG TCCTTTTTCG CCTCCCTGC 3180
CACCTCCCCG TGTTCACAG TCAGCTTTCG TGCAAGAAGA AATCCTGGTT AAAAAGTCT 3240
TTTGATTGG GTGAGGAGT GAATTTGGGG TGGGAGGATG GATGCAACTG AAGCAGAGT 3300
80 TGGGTGCCA GATGTGCGCT ATTAGATGTT TCTCTGATAA TGTCCTCAAT CATACAGGG 3360
AGACTGSCAT TGAAGAGAAC TCAGGTGGAG GCTTGAGAAG GCGCAAGGG CCCTGACCT 3420
GCCTGGCTTC CTAGCTTGC CCCTCAGCTT TGCAAGAGAG CACCTAGGC CCACTGAC 3480
CGCATGGGTG TGAGCCAGCT TGAGAACACT AACTACTCAA TAAAGCGAA GGTGGACAAA 3540
AAAAAAAAA AAAAA

Seq ID NO: 73 Protein sequence:
Protein Accession #: AAC51128.1

1 11 21 31 41 51
5 MKTSPRRPLI LKRRRLPLPV QNAPSETSEE EPKRSPAQQE SNQAEASKEV AESNSCKPFA 60
GIKI INEPTM PNTQVVAIPN NANIHISIITA LTAKGKESGS SGPKNFILIS CGGAPTQPPG 120
LRPQTQTSYD AKRTEVTLET LGPKPAARDV NLPKPPGALC EQKRETCADG EAAGCTINNS 180
LSNTQWLKRM SSDGLGSRSI QQEMEKEKENC HLEQRQVKVE EPSRPSASWQ NSVSERPPYS 240
YMAIMQFAIN STERKMTLK DIYTWIEDHF PYFKHIAKPG WWSIRHNLS LHDMPVRETS 300
10 ANGKVSFWTI HPSANRYLTL DQVFKPLDPG SPQLPEHLES QQKRFNPELR RNMTIKTELP 360
LGARRRMKPL LPRVSSYLVP IQFFVNQSLV LQPSVKVPLP LAASLMSSSEL ARHSKRVRIA 420
PKVFGQVVF GYMSKFFSGD LRDFGTPITS LFNFIPLCLS VLLAEEGIAP LSSAGPGKEE 480
KLLPGEQFSP LLPVQTIKEE EIQPGEEPMH LARPIKVES PLEEWSPAP SFKESSHWS 540
EDSSQSPTPR PKKSYSGLRS PTRCVSEMLV IQHRERRERS RSRRKQHLLP PCVDEPELLP 600
15 SEGPSTSRWA AELPPADSS DPASQLSYSQ EVGGPFKTP I KETLPISSSTP SKSVLPRTPE 660
SWRLTPFAKV GGLDFSPVQT SQGASDPLPD PLGLMDLSTT PLQSAPPLES PQRLLSSEPL 720
DLISVPFGNS SPSPIDVPEK GSPEPQVSGL AANRSLTEGL VLDTMNDSLS KILLDISFPG 780
LDEDPLGPDN INWSQPIPEL Q

20 Seq ID NO: 74 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 111-416

25 1 11 21 31 41 51
GGGAAGAGCC AGGCTGAGCC TTATAAAGGA CTGCTCTTTG TCCAAACACA CACATCTCAC 60
TCATCCTTCT ACTCGTGAGC CTTCCTCAGCT CTGGCTTTT GAAAGCAAAG ATGAGCAACA 120
CTCAAGCTGA GAGGTCCATA ATAGGCATGA TCGACATGTT TCACAAATAC ACCGACGCTG 180
ATGACAAGAT TGAGAAGCCA AGCCTGCTGA CGATGATGAA GGAGAACTTC CCCAACTTCC 240
30 TTAGTGCTGT TGACAAAAG GGCACAAAT ACCTCGCCGA TGTCTTTGAG AAAAAGGACA 300
AGAATGAGGA TAAGAAGATT GATTTTCTG AGTTTCTGTC CTGCTGGGA GACATAGCCA 360
CAGACTACCA CAAGCAGAGC CATGGAGCAG CGCCCTGTTC CGGGGGCAGC CAGTGACCCA 420
GCCCCACCAA TGGGCTTCCA GAGACCCAG GAACAATAAA ATGCTCTCTC CCACAGCA

35 Seq ID NO: 75 Protein sequence:
Protein Accession #: Eos sequence

40 1 11 21 31 41 51
MSNTQAERSI IGMIDMFHYK TRDDKIEKP SLLTMKENF PNFLSACDKK GTNYLADVFE 60
KDKKNEDKKI DFSEFLSLLG DIATDYHKQS HGAAPCSGGS Q

45 Seq ID NO: 76 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 111-416

50 1 11 21 31 41 51
GGGAAGAGCC AGGCTGAGCC TTATAAAGGA CTGCTCTTTG TCCAAACACA CACATCTCAC 60
TCATCCTTCT ACTCGTGACA CTTCCTCAGTT CTGGCTTTT GAAAGCAAAG ATGAGCAACA 120
CTCAAGCTGA GAGGTCCATA ATAGGCATGA TCGACATGTT TCACAAATAC ACCGACGCTG 180
ATGGCAAGAT TGAGAAGCCA AGCCTGCTGA CGATGATGAA GGAGAACTTC CCCAATTTC 240
TCAGTGCTGT TGACAAAAG GGCATACATT ACCTCGCCAC TGTCTTTGAG AAAAAGGACA 300
AGAATGAGGA TAAGAAGATT GATTTTCTG AGTTTCTGTC CTGCTGGGA GACATAGCCA 360
55 CAGACTACCA CAAGCAGAGC CATGGAGCGG CGCCCTGTTC TGGGGGAAGC CAGTGATCCA 420
GCCCCACCAA GGGGCTTCCA GAGACCCAG GAACAATAAG TGTCTCTCTC CACCAGA

60 Seq ID NO: 77 Protein sequence:
Protein Accession #: XP_048124.1

65 1 11 21 31 41 51
MSNTQAERSI IGMIDMFHYK TGRDGKIEKP SLLTMKENF PNFLSACDKK GIHYLATVFE 60
KDKKNEDKKI DFSEFLSLLG DIAADYHKQS HGAAPCSGGS Q

70 Seq ID NO: 78 DNA sequence
Nucleic Acid Accession #: Z73678.1
Coding sequence: 253-2433

75 1 11 21 31 41 51
GGGGTGGTGC AGGGCAGGGG TGGTATATCC TGTCTGACGG AGGGCGGGCC TCGCCAGTGC 60
CAGAGAGGGA CGAACCAGGG TGGGAAGCGCC AGGAGCAGCT GCAGGGAGCC CTCACGCGGA 120
CCTCGCACTC TATGGCCGTA GGGAGCGGCT GAGAGCGAGA AGAGCACGCT CCTGCCCGCC 180
CGCTGCACCG CACCTCGGCT CGCCTCTCTG CTCTCTAGG CCCCAGCGCG GCGGCAACCG 240
CCTCCGCGCA CCAATGAACA CTGCGCGCTC AAGACCGGCT TGGCGTACGA ATGCTTCCAG 300
GACCAGGACA ACTCCACGTT GGCTTTGCCG TCGGACCAAA AGATGAAAAC AGGCAGCTCT 360
GGCAGGAGC GCGTGCAGGA GCAGGTGATG ATGACCGTCA AGGGGAGAAA GTCCAAGTCT 420
80 TCCAGTCTGT CCACCTCGAT CCACTCCAAT CGAGGTTCCA TGTATGATGG CTTGGCTGAC 480
AATTACAAC ATGGGACAC CAGCAGGAGC AGCTACTACT CCAAGTTCCA GGCAGGGAAT 540
GGCTCATGGG GATATCCGAT CTACAATGGA ACCCTCAAGC GGGAGCTCGA CAACAGGGCC 600
TTCAGCTCTT ACAGCCAGAT GGAGAAGTGG AGCCGGCACT ACCCCCGGGG CAGCTGTAAC 660
ACCACCGCG CAGGACAGGA CATCTGCTTC ATGCAGAAAA TCAAGGCGAG CCGCAGTGAG 720
85 CCGACCTCT ACTGTGACCC ACGGGGCACC CTGCGCAAGG GCACGCTGGG CAGCAAGGGC 780
CAGAGACCA CCAGAACCG CTACAGCTTT TACAGCACT GCAGTGGTCA GAAGGCCATA 840
AAGAAGTGCC CTGTGCGCCC GCCCTCTGTT GCCTCAAGC AGGACCTGT GTATATCCCG 900

	CCCATCTCCT	GCAACAAGGA	CCTGTCTCTT	GGCCACTCTA	GGGCCAGTTC	CAAGATCTGC	960
	AGTGAGGACA	TGAGTGCAG	TGGGCTGACC	ATCCCCAAGG	CTGTGCAGTA	CCTGAGCTCC	1020
	CAGGATGAGA	AGTACCAGGC	CATTGGGGCC	TATTACATCC	AGCATACCTG	CTTCCAGGAT	1080
	GAATCTGCCA	AGCAACAGGT	CTATCAGCTG	GGAGGCACTT	GCAAGCTGGT	GGACCTCCTC	1140
5	CGCAGCCCCA	ACCAGAACGT	CCAGCAGGCC	CGCGCAGGGG	COCTGCGCAA	CCTGGTGTTC	1200
	AGGAGCACCA	CCAACAAGCT	GGAGACCCGG	AGGCAGAATG	GGATCCGCGA	GGCAGTCAGC	1260
	CTCCCTGAGGA	GAACCGGAA	CGCCGAGATC	CAGAAGCAGC	TGACTGGGCT	GCTCTGGAAC	1320
	CTGTCTTCCA	CTGACGAGCT	GAAGGAGGAA	CTCATTGCGG	ACGCCCTGCC	TGTTCTGGCC	1380
	GACCGCGTCA	TCATTCCCTT	CTCTGGCTGG	TGCGATGGCA	ATAGCAACAT	GTCCCGGAA	1440
10	GTGGTGGACC	CTGAGGCTTT	CTTCAATGCC	ACAGGCTGCT	TGAGGAACCT	GAGCTCGGCC	1500
	GATGACAGGC	GCCAGACCAT	GGTAACTAC	TCAGGGCTCA	TTGATTCCCT	CATGGCCTAT	1560
	GTCCAGAACT	GTGTAGCGGC	CAGCCGCTGT	GACGACAAGT	CTGTGGAAAA	CTGCATGTGT	1620
	GTTCGTGACA	ACCTCTCCTA	CGCCTGGAC	GCCGAGGTGC	CCACCCGCTA	CGCCAGCTG	1680
	GAGTATAACG	CCCGCAACGC	CTACACCGAG	AAGTCTCTCA	CTGGCTGCTT	CAGCAACAAG	1740
15	AGCGACAAGA	TGATGAACAA	CAACTATGAC	TGCCCCCTGC	CTGAGGAAGA	GACCAACCCC	1800
	AAGGCGAGCG	GCTGGTGTGA	CCATTAGAT	GCCATCCGCA	CCTACCTGAA	CCTCATGGGC	1860
	AAGAGCAAGA	AAGATGCTAC	CCTGGAGGCC	TGTGCTGGTG	COCTGCAGAA	CCTGACAGCC	1920
	AGCAAGGGGC	TGATGTCCAG	TGGCATGAGC	CAGTTGATTG	GGCTGAAGGA	AAAGGGCCTG	1980
	CCACAATATG	CCCGCTCTCT	GCAATCTGGC	AACCTCTGAT	TGGTGGGTC	CGAGCCTCC	2040
20	CTCCTGAGCA	ACATGTCCCG	CCACCCTCTG	CTGCACAGAG	TGATGGGAA	CCAGGTGTTC	2100
	CCGGAGGTGA	CCAGCTCTCT	CACCAGCCAC	ACTGGCAATA	CCAGCAACTC	CGAAGACATC	2160
	TGTCTCTCGG	CCTGCTACAC	TGTGAGGAAC	CTGATGGCCT	CGCAGCCACA	ACTGGCCAA	2220
	CAGTACTTGC	CAGCAGCATG	GCTCAACAAC	ATCATCAACC	TGTGCCGAAG	CAGTGCCTCA	2280
	CCCAAGGCCG	CAGAAGCTGC	CCGGCTTCTC	CTGTCTGACA	TGTGTCAG	CAAGGAAGTC	2340
25	CAGGGTGTCC	TCAGACAGCA	AGGTTTCGAT	AGGAACATGC	TGGGAACCTT	AGCTGGGGCC	2400
	AAACGCTCTA	GGAACCTCAC	CTCCGATTTC	TAAGAAGAGA	CTGTCCAAGC	AAGTTAGGCT	2460
	TGACGGAAGA	TATGACCCAG	CTGAGAAGCC	CTCAGGCCTC	GCTGGATGGG	GTTTCTGTGC	2520
	CATCCTGTGC	AGTATTGGGG	AAAGTTTACA	AGAAACTGAG	AAGAAACCTA	AAAACCTGTC	2580
	ATAGTGGAAA	GATTTTATGA	TTTTTTTTTT	CCTTGGGGAA	ACTGGCAGGC	AATGGGGGTT	2640
30	AGGGAGGTGG	GGGGGGGGGG	GGCTTTCTTG	AGTTAAAGGG	GCTTATATGT	GATGTCAATA	2700
	TTTCTTCTCT	TGAGAAATGG	TATATATATG	TGTCTAATGT	AAGTGTGTGC	ATGCATGTGC	2760
	GGGTGCTATG	GTGTGTGTGT	GAGTGTCTTA	AAGCATAACC	ACAACTGCA	AAAAGCTAGG	2820
	TAAGCTATT	TGTTGACGT	CATAAGGTGG	TGAAAAGGAC	TCTCTGTGT	TTCTTACTCA	2880
	TAGGCAAGGA	CAACATGTGC	TTTTTGGTGA	GCTGCTCATA	ATTCCTGAAA	TGTGTGGTGC	2940
35	CAGGGCAAGG	GGGCGATCAC	TGCAGTCAGG	COCTCAGAGG	AGTCTGTCAG	GCTTCTTACC	3000
	AGTGGTCTCC	AAGGGTGCAG	GAGTAACTGG	GGCTGGGCCA	GCCTCCCCCC	TTACAAGGCT	3060
	GCTTTCCACG	AAGGGAGGTC	TGGTGTATCT	CATGGGAGAA	TCTGGGGTGT	CTGTAGTGTG	3120
	ACCCCTCCAG	CAGCCCAACA	AGGACTGAGG	TTGGGTAGGT	GTGAGGTTC	AGAGGACAGC	3180
	AGGACACTCT	CCCATACTTT	GCCAAATGAG	GCCTGCTCAG	AGGAGTAGGA	GCTGAAAGAT	3240
40	GGTGCTCTCC	ACCTCTCTGG	GCTGTGTGCC	CATCAGAGCA	GGCTCAGCCT	GCAAAGGCC	3300
	TGCACTTACA	GGCTGTGTA	TCTACTTGT	GCAGGAGAAA	GAAGGTAAAA	AATGATTTTT	3360
	TTAAGAAAGG	CTATTTTATT	GCAGCTCTTT	CCCAAGAGCT	GTTCTGGGAA	TGGCTGGTCT	3420
	TCATATTCCC	AGTGGAGAGG	GGAAACAAGT	GGGCTGGGCA	TATACCTATT	CCGGCTTCTA	3480
	GTGGGATGGA	GTGGGGTAT	AGAAATTAAC	CAGGAAGATG	TTTCCACCAA	GCCTGCTGTG	3540
45	AGTCAATTGA	GGGAGTGT	GGGTCCCAGG	AGACTTGGAC	GGGGGAGT	TGGGTAGACT	3600
	AGGAAAGGAA	AGTGCCATAT	CAGGGTACCG	GTACCGGCAA	GCTCACATCT	CAGCCAGGGG	3660
	CCATGCCCCA	CTTCCCCTGA	CCCCAGCTGT	CTTGTCTCCA	CTCTGTGAAA	CCCCAGGGG	3720
	ATGTGATAAA	AGGGGCTATT	AGGGGTATCA	GCCACGTCGA	GCCCCCAGAC	TCTGTGCACT	3780
50	TCAGACCAGC	AGCAGCAGGA	GGGCTCCCGA	GGGCTTATG	AGAAAACCTG	TGTGGACATC	3840
	CCTTGGTGTA	CACATAAGACA	GAGCAGAGCC	CAGCGCTCCC	AAGCCTTCTT	CCTTCCAGCT	3900
	TCTACCTCCA	TGCTAGCATT	GCTGGTGTTA	GAGAGGAATT	AACCTTCTGG	TCTGTGCCCT	3960
	TCTCTAGAAG	AATATAAGAT	GCTCCTCTCT	CTCACCCCTT	CTCAGCCTCC	TCCCAAGTCT	4020
	TCCTCTCTG	CACCAACCCC	GAGTCCAAAC	CCACCTCTTG	CCCCAGCATT	CAGGCTGGAA	4080
	AACACTGATG	TGCACTCAGT	ATGACAACTG	AGATGGGGGA	AGCCAGACAT	GTGAGGACGC	4140
55	TGTCTCTCGA	GAGGTGTCCC	CGGCTGTTAG	CCAGCTGTGC	TGTGGTGTCT	TGGGTCTGTC	4200
	ATACCTCTCC	TGCTCTCTGT	TCACACTGGG	AGGCCCACTC	CTGGCTCACC	TCTCCTCTCT	4260
	AGGGACCCAC	GTGGAGACCT	GGATCCCTGG	ACTGTCTCGG	GCAATAGTTT	CAGGGGCTCT	4320
	CTTGTGTGTC	CTCAGAAACC	AGAGGAATTC	TTCTCTTAAA	AAATACGTAT	GGCATACCAA	4380
	TCTGTGCGGG	GCACTGTCTT	AAGCACTTAG	ACTACATCAG	GGAAGAACAC	AGACCACATC	4440
60	CCCGTCTCCA	TGCGGCTTAT	GTTTTCTGGA	GGAAAGTGGG	GACACAAGTC	CTTGGCTTTA	4500
	GGGCTCCCCC	GGCTGGGGGG	TGTGCAGTCC	GGTCAGGGCG	GGAGGGGAAA	TGCACCGCTG	4560
	CATGTGAACC	TTACCCAGCC	AGGCGGATGC	COCTTCCCTT	TAGCACTACC	CTGGCTCTCT	4620
	GCATCCCTCT	GCCTCAGTGT	CCTCCACCTT	TCAAAGAATG	AAGAGCCCCA	TGGGCCCAGC	4680
	CCCTGCCCTG	GGAACCAAGC	AGCCTTCCAG	ACCTCAGGGG	CTGAGGCAGA	CTATTAGGGC	4740
65	AGGGGTGACT	TTGGTGACAC	TGCCCATTC	CTCTCAGGCC	AGCTCAGGTC	ACCCGGGCTT	4800
	CTGACCCAGG	CCTGTCACTT	TGAGAGGGGC	AAAACCTGAGA	GGGGCTTTTC	CTAGAGAAAG	4860
	AGAACAAGGA	GCTTGCCAGG	CTTCATGTAG	CCGACACACG	TCTCAGGATT	TTAAGTCCAC	4920
	ATTGGCCTCA	CATAGCCTTA	GGCCAATGCC	CAAAATAAGG	AGTTCCAAAT	TGGGGCCAAA	4980
	TGAGGAAGGA	CACAGACTCT	GCCCTGGGAT	CTCCTGTGCT	AGCGGCCAAT	GACAAATCCA	5040
70	GTCAATTGGCC	ACCAGCCACC	TCTGCAGTGG	GGACCACACT	AGCAGCCCTG	ACTCCACACT	5100
	CCTCCTGGGG	ACCCAAGAGG	CAGTGTGTCT	GTCTGCGTGT	CCACCTTGGG	ATCTGGCTGA	5160
	ACTGGCTGGG	AGGACCAAGA	CTGCGGCTGG	GGTGGGCAGG	GAAGGGAAGC	CGGGGGCTGC	5220
	TGTGAGGGAT	CTTGAAGCTT	COCTGTAGCC	CACCTTCCCC	TGTCTTCATG	TTTGTAGAGG	5280
75	AACTTGTGTC	CGGCCAGGCC	CAGTTTCTCT	GTGTGATACA	CTAATGTATT	TGCTTTTTTT	5340
	GGAAATAGAG	AAAATCAATA	AATTGCTAGT	GTTTCTTTGA	AAAAAATA		

Seq ID NO: 79 Protein sequence:
Protein Accession #: CAA98022.1

80	1	11	21	31	41	51	
	MNHSPLKTA	AYECFQDQDN	STLALPSDQK	MKTGTSGRQR	VQEQVMVTVK	RQKSKSSQSS	60
	TLSHSNRSGM	YDGLADNYYN	GTTSRSSYYN	KFOAGNGSWG	YPIYNGTLKR	EPDNRFRSSY	120
	SQMNENWSRHY	PRGSCNTTGA	GSDICFMQXI	KASRSEPDLY	CDPRGTILRXG	TLGSKGQRTT	180
85	QNRYSFYSTC	SGQKAIKICP	VRPPSCASKQ	DPVYIPIISC	NKDLSPFHRS	ASSKICSEDI	240
	ECESGLTIPKA	VQYLSSQDEK	YQAIGAYYIQ	HTCFQDESAK	QQVYQLGGIC	KLVDLLRSPN	300
	QNVQAAAGA	LRNVLFSTT	NKLETRRQNG	TREAVSLLRR	TGNABIQKQL	TGLLWNLSST	360

WO 02/086443

PCT/US02/12476

DELKEELIAD ALPVLADRV I PPSGWCDGN SNMSREVVDV EVFFNATGCL RNLSSADAGR 420
 QTMRYNSGLI DSLMAYVQNC VAASRCDDKS VENCVCVLEN LSYRLDAEVP TRYRQLEYNA 480
 RNAYTEKSTT GCFSNKSDKM MNNYDCPLP EESTNPKGSG WLYHSDAIRT YLNLMGKSKK 540
 DATLEACAGA LQNLTASKGL MSSGMSQLIG LKEKGLPQIA RLLQSGNSDV VRSGASLLSN 600
 MSREPLLEHV MGNQVFPEVT RLLTSHGTNT SNEEDILSSA CYTVRNLMAS QPQLAKQYFS 660
 SSMNNIINL CRSSASPKAA EAARLLLSDM WSSKELQGV L RQQGFDRLML GTLAGANSLR 720
 NFTSRF

Seq ID NO: 80 DNA sequence
 Nucleic Acid Accession #: NM_006516.1
 Coding sequence: 180-1658

1 11 21 31 41 51
 TAGTCGCGGG TCCCGGAGTG AGCACGCCAG GGAGCAGGAG ACCAAACGAC GGGGGTCGGA 60
 GTCAGAGTCG CAGTGGGAGT CCCCGGACCG GAGCAGCAGC CTGAGCGGGA GAGCGCCGCT 120
 CGCACGCGCG TCGCGACCCG CGTACCCGCG GACGACGAGC CCACGAGCGC AGCGCTGCCA 180
 TGGAGCGCGC CAGCAAGAAG CTGACGGGTC GCCTCATGCT GGCTGTGGGA GGAGCAGTGC 240
 TTGGCTCCCT CGAGTTTGGC TACAACACTG GAGTCATCAA TGCCCCCAG AAGGTGATCG 300
 AGGAGTTCTA CAACACAGCA TGGGTCCACC GCTATGGGGA GAGCATCCTG CCCACCAACG 360
 TCACCAACGT CTGGTCCCTC TCAGTGGCCA TCTTTTCTGT TGGGGGCATG ATTGGCTCCT 420
 TCTCTGTGGG CCTTTTCGTT AACCGCTTTG GCGCGCGGAA TTCAATGCTG ATGATGAAAC 480
 TGCTGGCCTT CGTGTCCGCC GTGCTCATGG GCTTCTCGAA ACTGGGCAAG TCCTTTGAGA 540
 TGCTGATCCT GGGCGCGCTC ATCATCGGTG TGTACTGCGG CCTGACCACA GGCTTCGTGC 600
 CCATGTATGT GGGTGAAGTG TCACCCACAG CCTTTCGTGG GSCCCTGGGC ACCCTGCACC 660
 AGCTGGGCAT CGTGTGGGCG ATCCTCATCG CCCAGGTGTT CGSCCTGGAC TCCATCATGG 720
 GCAACAGGGA CCGTGTGGCC CTGCTGCTGA GCATCATCTT CATCCCGGCC CTGCTGCAGT 780
 GCATCGTGCT GCCTTCTGCG CCGGAGAGTC CCGCTTCTCT GCTCATCAAC CGCAACGAGG 840
 AGAACCGGGC CAAGAGTGTG CTAAGAAGC TGCGCGGGAC AGCTGACGTG ACCCATGACC 900
 TGCAGGAGAT GAAGGAAGC AGTCGGCAGA TGATGCGGGA GAAGAAGGTC ACCATCTGG 960
 AGCTGTTCGG TCCCCCGGCC TACCGGACAG CCATCCTCAT CGCTGTGGTG CTGCACTGCT 1020
 CCCAGCAGCT GTCTGCGATC AACGCTGTCT TCTATTACTC CACGAGCATC TTGAGAGAAG 1080
 CGGGGGTGCA CGAGCTGCTG TATGCCACCA TTGGCTCCGG TATGCTCAAC ACGGCTTCA 1140
 CTGTGCTGTC GCTGTTTGTG GTGAGCGGAG CAGGCGCGCG GACCCTGCAC CTCATAGGCC 1200
 TGCTGGCATG GGGCGGTGTT GCCATACTCA TGACCATCGC GCTAGCACTG CTGAGCAGC 1260
 TACCTTGGAT GTCTATCTG AGCATCGTGG CCATCTTTGG CTTTGTGGCC TTCTTTGAAG 1320
 TGGGTCTCGG CCCCATCCCA TGGTTCATCG TGGCTGAAGT CTTAGCCAGG GGTCCACGTC 1380
 CAGCTGCCAT TGCCGTTCTG GGCTTCTCCA ACTGGACCTC AAATTTTATT GTGGGCATGT 1440
 GCTTCCAGTA TGTGAGGCAA CTGTGTGTCG CTAAGTCTCT CATCATCTTC ACTGTGCTCC 1500
 TGGTCTCTGT TCTACTCTTC ACCTACTTCA AAGTTCCTGA GACTAAAGGC CGGACCTTGG 1560
 ATGAGATGCG TTCCGCTTTC CGGCAGGGGG GAGCCAGCCA AAGTGATAAG ACACCCGAGG 1620
 AGCTCTTCCA TCCCTCTGGG GCTGATTCCC AAGTGTGAGT CGCCCCAGAT CACGAGCCCG 1680
 GCCTGCTCCC AGCAGCCTTC AGGATCTCTC AGGAGCAGAG GCAGCTGGAT GAGACTTCCA 1740
 AACCTGACAG ATGTGAGCGC AGCGGGGCTT GGGGCTCTCT TCTCCAGCCA GCAATGATGT 1800
 CCAGAAGAAAT ATTACAGGAT TAACGGCTCC AGGATTTTAA CAAAGCAAG ACTGTTGCTC 1860
 AAATCTATTC AGACAAGCAA CAGGTTTTAT AATTTTTTTA TTAAGTATTG TGTATTTTT 1920
 ATATCAGCCT GAGTCTCTCG TGCCACATC CCAGGCTTCA CCTGAATGG TTCCATGCTG 1980
 GAGGGTGGAG ACTAAGCCCT GTGAGACAC TTGCTTCTCT CACCCAGCTA ATCTGTAGGG 2040
 CTGACCTTAT GTCTTAAGGA CACACTAATC GAAGTATGAA CTACAAAGCT TCTATCCAG 2100
 GAGGTGGCTA TGCCACCCG TTCTGCTGCG CTGATCTCC CCACTTAGG GTTACGGCTC 2160
 CATTAGGATT TGCCCTTCTC CATCTCTTCC TACCCACCA CTCAAATTAA TCTTTCTTTA 2220
 CCTGAGACCA GTTGGGAGCA CTGGAGTGCA GGGAGGAGAG GGGAGGGGCC AGTCTGGGCT 2280
 GCGGGGTTCT AGTCTCTCTT GCACTGAGGG CCACTATTAT ACCATGAGAA GAGGGCCTGT 2340
 GGGAGCCTGC AAATCACTG CTCAAGAAGA CATGGAGACT CTGCGCTGT TGTGTATAGA 2400
 TGCAAGATAT TTATATATAT TTTTGGTTGT CAATATTAAA TACAGACACT AAGTTATAGT 2460
 ATATCTGGAC AAGCCAACTT GTAAATACAC CACCTCACTC CTGTTACTTA CCTAAACAGA 2520
 TATAAATGGC TGGTTTTTAG AAACATGTTT TTGAATGCT TGTGGATTGA GGGTAGGAGG 2580
 TTTGGATGGG AGTGAAGTGG AAGTAAGTGG GGTGCAACG ACTGCAACGG CTAGACTTCT 2640
 GACTCAGGAT CCAGTCCCTT ACACGTACCT CTCATCAGTG TCCTCTGCT CAAAATCTG 2700
 TTTGATCCCT GTTACCCAGA GAATATATAC ATTCTTATC TTGACATTA AGGCATTTCT 2760
 ATCAGATATT TGATAGTTGG TGTCAAAAA AACACTAGTT TTGTGCCAGC CGTATGCTC 2820
 AGGCTTGAAA TGCATTATT TTGAATGTGA AGGGAA

Seq ID NO: 81 Protein sequence:
 Protein Accession #: NP_006507.1

1 11 21 31 41 51
 MEPSKKLTG RLMLAVGGAV LGSLOFGYNT GVINAPQKVI BEFYNTQVWH RYGESILPPT 60
 LTTLSLSVA IFSVGMIGS FSVGLFVNR GRNNSMLMMN LLAPVSAVLM GFSKLGKSE 120
 MLILGRFIIG VYCGLTGTFV PMYVGEVSPT AFRGALGTLH QLIGVVGILI AQVFGLDSIM 180
 GNRLWLPLL SIIPIPALLO CIVLPFCPE PRFLINRNE ENRAKSVLKK LRGTADVTHD 240
 LQEMKEESRQ MMREKVTIL ELFRSPAYRQ PILIAVVLQ SQQLSGINAV FYSTSTIFEK 300
 AGVOQPVYAT IGSIVNTAP TVVSLFVVER AGRRTLHLIG LAGMAGCAIL MTIALALLEQ 360
 LFWMSYLSIV AIFGFAFFE VGPPIPWFI VAELEFSGPR PAALAVAGFS NWTNFIIVGM 420
 CFQYVEQLCG PYVFIPTVL LVLFFIPTYF KVPETKGRTP DEIASGFRQG GASQSDKTP 480
 ELPHPLGADS QV

Seq ID NO: 82 DNA sequence
 Nucleic Acid Accession #: BC001291
 Coding sequence: 44-541

1 11 21 31 41 51
 GGGGGCGCGG CGCGCTGACC CTCCTTGGGC ACCGCTGGGG ACAGTGGGCG TGCTCGCCTT 60
 GCTGCTGGTC GTGGCCCTAC CGCGGGTGTG GACAGACGCC AACCTGACTG CGAGACAAAG 120

5 AGATCCAGAG GACTCCAGC GAACGGACGA GGGTGACAA AGAGTGTGGT GTCATGTTTG 180
 TGAGAGAGAA AACACTTTTC AGTGCCAGAA CCCAAGGAGG TGCAATGGA CAGAGCCATA 240
 CTGCGTTATA GCGCCGTGA AATATTTCC ACGTTTTTC ATGTTGCGA AGCAGTGCTC 300
 CGCTGGTTGT GCAGCGATGG AGAGACCCAA GCCAGAGGAG AAGCGTTTC TCCTGGAAGA 360
 GCGCATGCC TCTTTTACC TCAAGTGTG TAAATTCGC TACTGCAAT TAGAGGGGGC 420
 ACCTATCAAC TCATCAGTGT TCAAAGAATA TGCTGGGAGC ATGGGTGAGA GCTGTGGTGG 480
 GCTGTGGCTG GCCATCTCC TGCTGTGGC CTCCATGCA GCCGCGCTCA GCCTGTCTTG 540
 AGCCACGGGA CTGCCACAGA CTGAGCCTTC CGGAGCATGG ACTGCTCCA GACCGTTGTC 600
 10 ACCTGTGCA TAACTTGT TTTCTGTGA TTACCTCTTG GTTTGACTTC CCAGGCTCTT 660
 GGGATGGAG AGTGGGATC AGGTGCAGTT GGCTCTTAA CCTCAAGGGT TCTTTAACTC 720
 ACATTCAGAG GAAGTCCAGA TCTCCTGAGT AGTGATTTTG GTGACAAGTT TTTCTCTTGG 780
 AAATCAAACC TTGTAACAT TTTATTGCTG ATGGCCACTC TTTTCTTGA CTCCCTCTTG 840
 CCTCTGAGGG CTTCAGTATT GATGGGAGG GAGGCCAAG TACCCTCAT GGAGATATG 900
 TGCTGAGATG CTCCGACCT TCCAGTGAC GCAGGAACAC TGGGGGAGTC TGAATGATTG 960
 15 GGGTGAAGAC ATCCCTGGAG TGAAGGACTC CTCAGCATGG GGGGCGATGG GGCACAGTTT 1020
 AGGGCTGCC CCATTCCAGT GGTGGAGGGC CTGTGGATGG CTGCTTTCC TCAACCTTTC 1080
 CTACCAAGAT CCAGGAGGCA GAAGATAACT AATGTGTGTG AAGAACTTA GACTTCACCC 1140
 ACCAGCTGGC ACAGGTGCAC AGATTCTATA ATTCCACAC GTGTGTGTTT AACATCTGAA 1200
 20 ACTTAGGCCA AGTAGAGAGC ATCAGGGTAA ATGGCGTTCA TTTCTCTGTT AAGATGCAGC 1260
 CATCCATGGG GAGCTCAGAA ATCAGACTCA AAGTTCCACC AAAACAAT ACAGGGGAC 1320
 TTCAAAAGTT CAGAAAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA

Seq ID NO: 83 Protein sequence:
Protein Accession #: AAH01291

25 1 11 21 31 41 51
 MALLALLLVV ALPRVWTDAN LTARQRPED SQRTDEGDNV VWCHVCEREN TPECQNPRLR 60
 30 KWTPEYCVIA AVKIPRPFM VAKQCSAGCA AMERPKPEEK RPLLEPMFP FYLKCKKIRY 120
 CNLEGGPINS SVFKEYAGSM GESCGGLWLA ILLLLASIAA GLSL

Seq ID NO: 84 DNA sequence
Nucleic Acid Accession #: NM_022893.1
Coding sequence: 229-2726

35 1 11 21 31 41 51
 TTTTITTTT TTTTITGCTT AAAAAAAGC CATGACGGCT CTCCACAAT TCATCTTCCC 60
 40 TGCGCCATCT TTGTAATTAT TCTAATTAT TTTGGATGTC AAAAGGCACT GATGAAGATA 120
 TTTTCTCTGG AGTCTCTTTC TTTCTAACCC GGCTCTCCCG ATGTGAACCG AGCCGTCGTC 180
 CGCCCGCGGC CGCCCGCGCC GCGCGCGCCG CCGCGCCCGC AGCCCACTAT GTCTCGCGGC 240
 AAGCAAGCA AACCCGACCA CTTAAGCAAA CGGGAATTCT CGCCCGAGCC TCTTGAAGCC 300
 ATTCTTACAG ATGATGAACC AGACCAAGCG CCGTTGGGAG CTCAGAAGG GGATCATGAC 360
 45 CTCTCACTCT GTGGGCGATG CCAGATGAAC TTCCCATTTG GGGACATTCT TATTTTATATC 420
 GAGCACAAC GGAACAATG CAATGGCAGC CTCTGCTTAG AAAAGCTGT GGATAAGCCA 480
 CTTTCCCTTT CACCAATCGA GATGAAAAA GCATCCATC CCGTGGAGGT TGGCATCCAG 540
 GTCAAGCCAG AAGATGAAGA TTTTATATCA ACCTCATCTA GAAGAATTGG CCCCAACAC 600
 GAACACATAG CAGATAAAGT TCTGCACTGG AGGGGCTCT CTCCCTCTCG TTTGACATG 660
 GGAGCTCTAA TCCCAAGGCC TGGGATGAGT GCAGAATATG CCGCGCAGGG TATTGTGAAA 720
 50 GATGAGCCCA CGAGCTACAC ATGTACAAC TGCAACACAG CATTACACAG TGATGGTTT 780
 CTCTTGCAAC ACGCAGAGAA CACTCATGGA TTAAGAATCT ACTTAGAAG CGAACACGGA 840
 AGTCCCTGTA CCGCGCGGGT TGGTATCCCT TCAGGACTAG GTGCAGATG TCCTTCCCAG 900
 CCACCTCTCC ATGGGATTTA TATTGACAGC AATAACCCCT TTAACCTGCT AAGAATACCA 960
 55 GGATCAGTAT CGAGAGAGGC TTCCGCGCTG GCAGAAGGGC GCTTTCACCC CACTCCCCCC 1020
 CTGTTTAGTC GACCAACGAG ACATCACTTG GACCCCAACC GCATAGAGCG CTGGGGGGCG 1080
 GAAGAAATGG CCTTGGCCAC CCATCAACCC AGTGCCTTTG ACAGGGTGTCT GCGGTGAAT 1140
 CCAATGGCTA TGGAGCTTCC CGCCATGGAT TTCTCTAGGA GACTTAGAGA GCTGGCAGGG 1200
 AACAGCTCTA GCGCACCGCT GTCCCGAGGC CGGCCAGCGC CTATGCAAG GTTACTGCAA 1260
 60 CCATTCCAGC CAGGTAGCAA GCGGCCCTTC CTGGCGAGCG CCGCCCTCCC TCCTCTGCAA 1320
 TCAGCCCTCT CTCCCTCCCA GCGCCCGGTC AAGTCAAGT CATGCGAGTT CTGGGCAAG 1380
 ACGTTCAAT TTCAAGAGCA CCTGTGTGTG CACCGCGGCA GGCACACGG CGAGAAGCCC 1440
 TACAAGTGA ACCTGTGCGA CCACGCGTGC ACCAGGCCA GCAAGCTGAA GCGCCACATG 1500
 AAGAGCACA TGCACAATC GTCCCATATG ACGGTCAAGT CCGACGACGG TCTCTCCACC 1560
 65 GCCAGCTCCC CGGAACCGCG CACCAAGGAG TTGGTGGGCA GCGCCAGCAG CGCGCTCAAG 1620
 TCCTGTGGTG CCAAGTTCAA GAGCGAGAAC GACCCCAACC TGATCCCGGA GAACGGGGAC 1680
 GAGGAGGAG AGGAGGACGA CGAGGAAGAG GAAGAAGAGG AGGAAGAGGA GGAGGAGGAG 1740
 CTGACGGAGA GCGAGAGGGT GGAATACGGC TTGGGCTGTA GCCTGAGGC GCGCGCCAC 1800
 CACGAGAAC GCTCGCGGGG CGCGGTGCTG GCGTGGGGCG ACGAGAGCG CGCCCTGCC 1860
 70 GACGTATGTC AGGGCATGGT GCTCAGCTCC ATGCAGCACT TCAGCGAGGC CTTCACACAG 1920
 GTCTTGGCG AGAAGCTATA GCGCGGCAC CTGGCGAGG CCGAGGGCCA CAGGGACACT 1980
 TGGAGCAAG ACTCGGTGGC CGCGAGTCTG GACCGCATAG ACGATGGCAC TGTATATGGC 2040
 CGCGGCTGCT CCGCGGGCGA GTCGGCTCTG GGGGGCTGT CCAAAAGCT GCTGTGGGTC 2100
 AGCCCGGCTA CGCTGAGCCC CTCTCTAAG CGCATCAAGC TCGAGAAGGA GTTGCACCTG 2160
 75 CCGCGGCTCA CGATGCCAA CAOGGAGAAC GTGTACTCGC AGTGGCTCGC CGGCTACGCG 2220
 GCTCTCAGGC AGCTCAAAG TCCCTTCTTT AGCTTGGAG ACTCCAGACA ATGCGCTTTT 2280
 GCTCTCTCGT CGGAGCACTC CTGGGAGAAC GGGAGCTTGC GCTTCTCCAC ACGCGCGGG 2340
 GAGCTGAGCG GAGGATCTC GGGGCGCAGC GGCACGGGAA GTGGAGGAG CAOGCCCCAT 2400
 ATTAGTGTCT CGGGCACGGG CAGGCCACAG TCAAAAGAGG GCAGACGAG CGACACTTGT 2460
 80 GAGTACTGTG GGAAGTCTT CAAGAACTGT AGCAATCTCA CTGTCCACAG GAGAAGCCAC 2520
 ACGGGCGAAA GGCCTTATAA ATGCGAGCTG TGCAACTATG CCGTGGCCA GAGTAGCAAG 2580
 CTCACGAGC ACATGAAAC GCATGGCCAG GTGGGGAAGG ACGTTTACAA ATGTGAAATT 2640
 TGTAGATGC CTTTATAGCT GTACAGTACC CTGGAGAAAC ACATGAAAA ATGGCACAGT 2700
 GATGAGTGT TGAATAATGA TATAAAACT GAATAGAGT ATATTAATAC CCTCCCTCA 2760
 85 CTCCACCTG ACACCCCTT TTTCACTT CCTTTTCCCC ATGCGCTCC AGCCCACTC 2820
 CCTGTAGGAT TTTTCTTAG TCCCATGTGA TTTAAACAAA CAAACAAACA AACAGAGTA 2880
 ACGAAGCTAA GAATATGAGA GTGCTTGTGA CCAGCACACC TGTTTTCTTT TTTTCTCTT 2940
 TCTTTTTTTC TTTTCTCTT TTTTTTTTTC TCCTTTATGT TCTACCGTT TGAATGCATG 3000

	ATCTGTATGG	GGCAATACTA	TTCATTTTA	GGCAAACTTT	GAGCCTTTCT	CTGTGCAAT	3060
	AATTTACATG	TTGTGTATGT	TTTTTTTAA	ACITAGACAG	CATGTATGGT	ATGTTATGGC	3120
	TATTTTAAAT	TGTCCTTAAT	TCGTTGCTGA	GCAACATGT	TGCTGTTTCC	AGTTCOGTTC	3180
	TGAGAGAAAA	AGAGAGAGAG	AGAGAAAAAG	ACCATGCTGC	ATACATTCTG	TAATACATAT	3240
5	CATGTACAGT	TTTATTTTAT	AACGTGAGGA	GGAAAAACAG	TCCTTGGAAT	AACCTCTCAT	3300
	AGACAGAATA	GATAGCACTG	AAAAAAATC	TCTATGAGCT	AAATGTCTGT	CTCTAAAGGG	3360
	TTAAATGTAT	CAATTGGAAA	GGAGAAAAAA	AGGCCTTGAA	TTGACAAAT	AACAGAAAAA	3420
	CAGAACAGT	TTATTCTATC	ATTGTTT	AAATATGAG	TGCTTGGAT	CTATTAAAC	3480
	CACATCGATG	GTTCTTTCTA	CTTGTTATAA	ACTTGTAGCT	TAATTCAGCA	TTGGGTGAGG	3540
10	TAATAAACCT	TAGGAACCTAG	CATATAATTC	TATATTGTAT	TTCTCACAAC	AATGGCTACC	3600
	TAAAAAGATG	ACCCATTATG	TCCTAGTTAA	TCATCATTTT	TCCTTAGTGT	TAATTTTATA	3660
	AACAAAACCTG	ATTATACCAG	TATAAAAGCT	ACTTGTCTCC	TGGTGAGAGC	TTAAAAAGAA	3720
	TGGGCTGTTT	TGCCCAAGT	TTTATTTTT	TTAAACAATG	ATTAAATTGA	ATGTGTAATG	3780
	TGCAAAAGCC	CTGGAAAGCA	ATTAAATACA	CTAGTAAGGA	GTTTATTATA	TGAAGATATT	3840
15	TGCTTTAATA	ATGCTTTTTT	AAAAATCTG	GCACCAAAAG	AAATAGATCC	AGATCTACTT	3900
	GGTTGTCAAG	TGGACAATCA	AATGATAAAC	TTTAAGACCT	TGTATACCAT	ATTGAAAGGA	3960
	AGAGGCTGAC	AATAAGGTTT	GACAGAGGGG	AACAGAAAGAA	AATAATATGA	TTTATTAGCA	4020
	CAACGTGGTA	CTATTGCTCA	TTTAAAACTA	GAACAGGTAT	ATAAGCTAAT	ATTGATACAA	4080
	TGATGATTAA	CTATGAATTC	TTAAGACTTG	CATTTAAATG	TGACATTCCT	AAAAAAAGAA	4140
20	GAGAAAGAA	TTTAAGAGTA	GCAGTATATA	TGCTGTGCT	CCCTAAAGT	TGACTTCAT	4200
	TTCTTTTCCA	TACACTGTGT	GCTATTGTG	TTAACATGGA	AGAGGATTCA	TTGTTTTTAT	4260
	TTTTTTTTTT	TTAATTTTTT	CTTTTTTAT	AAGCTAGCAT	CTGCCCCAGT	TGGTGTTCAT	4320
	ATAGCACTTG	ACTCTGCTG	TGATATCTGT	ATCTTTCTC	TAATCAGAGA	TACAGAGGTT	4380
	GAGTATAAAA	TAAACCTGCT	CAGATAGGAC	AATTAAGTGC	ACTGTACAAT	TTTCCAGATT	4440
25	TACAGGTCTA	TACTTAAGGG	AAAAGTTGCA	AGAATGCTGA	AAAAAAATTG	AACACAATCT	4500
	CATTGAGGAG	CATTTTTTAA	AAACTAAAAA	AAAAAAACT	TTGCCAGCCA	TTTACTTGAC	4560
	TATTGAGCTT	ACTTACTTGG	ACGCAACATT	GCAAGCGCTG	TGAATGGAAA	CAGAATACAC	4620
	TTAACATAGA	AATGAATGAT	TGCTTTGCT	TCTACAGTGC	AAGGATTTTT	TTGTACAAAA	4680
	CTTTTTTAAA	TATAAATGTT	AAGAAAAAAT	TTTTTTAAAA	AACACTTCAT	TATGTTTAGG	4740
30	GGGGAACCTG	ATTTTAGGTT	TCCATTGCT	TGGTGGTGT	ACAAGACTTG	TTATCCATT	4800
	AAAAATGGTA	GTGGAAATTC	TATGCCTTGG	ATACACACCG	CTCTTCAGGT	TGTAAAAAAA	4860
	AAAAACATAC	ATTGGGGAAA	GGTTTAAGAT	TATATAGTAC	TTAAATATAG	GAAATGCAC	4920
	ACTCATGTG	AATGGCGGTT	TAAAATACAT	TTATGGTCTT	TTTTCTGTAT	TTCTAGAAATG	4980
	GTATTTGAAT	TAAATGTTC	TCTAGTGTTA	GGCACTATAG	TATTTATATT	GAAGCTTGTA	5040
35	TTTTTAACCTG	TTGCTGTTC	TCCTTAAAGG	TATCAATGTA	CTTTTTTGG	TAGTGGAAAA	5100
	AAAAAAGACA	GGCTGCCACA	GTATATTTTT	TTAATTGGC	AGGATAATAT	AGTGCAAATT	5160
	ATTTGTATGC	TTCAAAAAA	AAAAAAGAG	AGAAACAAAA	AAGTGTGACA	TTACAGATGA	5220
	GAAAGCATAT	AATGGCGGTT	TGGGGGAGCC	TGCTAGAAATG	TCACATGGAT	GGCTGTCTA	5280
	GGGGTGTATC	ATATCTGTTT	TTGTCCTTT	TTCTGTCTGC	CATCTGTAT	GCAGTACTGC	5340
40	AAGCTAATAA	CGTGTGTTT	TTATGTAGTG	TGCTTTTTGT	CCCTTTCCTT	CTATCACCCCT	5400
	ACATTCCAGC	ACTTTACCTT	CATATGCAGT	AAAAGAAAGA	AAGAAAAAAA	AAGGAAAAAA	5460
	AAAAAAAAC	CAATGTTTTT	CAGTTTTTTT	CATTGCCAAA	AACTAAATGG	TGCTTTATAT	5520
	TTAGATTGGA	AAGAAATTTCA	TATGCAAAAGC	ATATTAAAGA	GAAAGCCCGC	TTTAGTCAAT	5580
	ACTTTTTTGT	AAATGGCAAT	GCAGAAATAT	TTGTTATGCG	CCTTTTCTAT	TCCTGTAATG	5640
45	AAAGCTGTTT	GTCTGAATCT	GAAATTTTAT	CTTTTACTAT	GGGAGTCACT	ATTTATTATT	5700
	GCCTTAGTGG	CCGTGTTCAA	ACAGAGGCAC	TTAATTTGAT	CTTTTATTTT	TCCTTTGTTT	5760
	TATTTTTTTT	TTTATTAGTA	TGACCAAAAG	TCATTACAAC	CTGGCTTTTT	ATTGTATTGT	5820
	TTCTGTGCT	TTGTTAAGTT	CTATTGGAAA	AACCACTGTC	TGTGTTTTTT	TGGCAGTTGT	5880
50	CTGCATTAA	CTGTTTATAC	ACCCATTTTG	TCCTTTTATT	GAAAAATAA	AAAAAATTAA	5940

Seq ID NO: 85 Protein sequence:
Protein Accession #: NP_075044.1

55	1	11	21	31	41	51	
	MSRRKQKGPQ	HLKSKREFSPE	PLEAILTDDE	PDHGPLGAPE	GOHDLTTCGQ	QOMNFFLGDI	60
	LIFIEHKRKQ	CNGSLCLEKA	VDKPPSPSPI	EMKKASNPVE	VGIQVTPEDD	DCLSTSSRRI	120
	CPKQEHIAKD	LLHWRGLSSP	RSAGHALIPT	PGMSAEYAPQ	GICKDEPSSY	TCTCKQPFNT	180
60	SAWFLIQAQ	NTGELIRIYLE	SEHGSPLTPR	VGIPLGLGAE	CPSPPLHGI	HIADNNPFLN	240
	LRIPGSVSRE	ASGLAEGFRP	PTPPLFSPPP	RHLDLPHRIE	RLGAEEMALA	THHPSAFDRV	300
	LRLNPMAMEP	PAMDFSRRLR	ELAGNTSSPP	LSPGRPSPMQ	RLLPQFPQGS	KPPFLATPPL	360
	PPLQSAPPSP	QPPVKSKECE	FCGKTFKFQS	NLVVHRRSHT	GEKPYKCNLC	DHACTQASKL	420
	KRHMKTMHK	SSPMTVKSDD	GLSTASSPEP	GTSDLVGSAS	SALKSVVAKF	KSENDPNLIP	480
65	ENGDEEEED	DEEEEBEED	EEELTESER	VDYFGLSLE	AARHHENSSR	GAUVGVGDES	540
	RALPDVMQGM	VLSSMQHFSE	APHQVLGEKH	KRGHLAEAEQ	HRDTCOEDSV	AGESDRIDDG	600
	TVNRRGCSPP	ESASGGLSKK	LLLGSPSSL	PFSKRILEK	EFDLPPTAMP	NTENVYSQWL	660
	AGYAAARQLK	DPFLSFGDSR	QSPFASSEH	SSENGSLRFS	TPPGELDGGI	SGRSGTSGGG	720
	STPHISGPGT	GRPSSKEGRR	SDTCBYCGKV	FKNCSNLTVH	RRSHTGERPY	KCELCNYACA	780
70	QSSKLTRMK	THGQVQKDVY	KCEICKMPFS	VYSTLEKRMK	KWHSRDLVNN	DIKTE	

Seq ID NO: 86 DNA sequence
Nucleic Acid Accession #: XM_035292.2
Coding sequence: 53-1576

75	1	11	21	31	41	51	
	GCTCGCTGGG	CGCGGCTCC	CGGGTGTCCC	AGGCCCGGCC	GGTGCGCAGA	GCATGGCGGG	60
	TGCGGGCCCC	AAGCGGCGCG	CGCTAGCGGC	GCCGGCGGCC	GAGGAGAAGG	AAGAGGCGCG	120
	GGAGAAGATG	CTGGCCGCCA	AGAGCGCGGA	CGGCTCGGCG	CCGGCAGGCG	AGGCGGAGGG	180
	CGTGACCCCTG	CAGCGGAACA	TCACGCTGCT	CAACGGCGTG	GCCATCATCG	TGGGACCAT	240
	TATCGGCTCG	GGCATCTTCG	TGACGCCAC	GGCGTGCTC	AAGGAGGAG	GCTCGCGGG	300
	GCTGGCGCTG	GTGGTGTGGG	CCGCGTGGG	CGTCTTCTCC	ATCGTGGGCG	CGCTCTGCTA	360
85	CGCGGAGCTC	GGCACCACCA	TCTCCAAATC	GGCGGCGGAC	TACGCTTACA	TGCTGGAGGT	420
	CTACGGCTCG	CTGCCGCTCT	TCCTCAAGCT	CTGGATCGAG	CTGCTCATCA	TCCGCGCTTC	480
	ATCGCAGTAC	ATCGTGGCCC	TGGTCTTGGC	CACCTACCTG	CTCAAGCCGC	TCTTCCCCAC	540

CTGCCCGGTG CCGAGGAGG CAGCCAAGCT CGTGGCTGCT CTCTGGGTGC TGCTGTCTAC 600
GGCCGTGAAC TGCTACAGCG TGAAGGCGGC CACCCGGGTC CAGGATGCCT TTGCCCGCGC 660
CAAGCTCTCTG GCCTGGGCC TGATCATCCT GCTGGGCTTC GTCCAGATCG GGAAGGGTGA 720
TGTTCTCAAT CTAGATCCCA ACTTCTCATT TGAAGGCACC AAACCTGGATG TGGGGAACAT 780
5 TGTTGTGCA TTATACAGCG CCTCTTTTGC CTATGGAGGA TGAATTACT TGAATTTCTG 840
CACAGAGGAA ATGATCAACC CCTACAGAAA CCTGCCCTGC GCCATCATCA TCTCCCTGCC 900
CATCGTGAGC CTGGTGATCG TGCTGACCAA CCTGGCTAC TCCACCACC TGTCACCGA 960
GCAGATGCTG TCGTCCGAGG CCGTGGCCGT GGACTTCGGG AACTATCAC TGGGCGTCAT 1020
10 GTCTGGATC ATCCCGTCT TCGTGGGCTT GTCTGCTTC GGCTCCGTCA ATGGGTCCCT 1080
GTTTCATCC TCCAGSCTCT TCTTCGTGGG GTCCCGGAA GGCCACCTGC CCTCCATCCT 1140
CTCCATGATC CACCCACAGC TCCTCACCCC CGTGGCTGCC CTGTGTTC CA GTGTGTGAT 1200
GACGCTGCTC TACGCTTCT CCAAGGACAT CTCTCCGTC ATCAACTTCT TCAGCTTCTT 1260
CAACTGGCTC TGGTGGGCC TGGCCATCAT CGGCATGATC TGGCTGGCC ACAGAAAGCC 1320
TGAGCTTAG CCGCCATCA AGGTGAACCT GGCCCTGCCT GTGTCTTCA TCCTGGCCTG 1380
15 CCTCTTCTC ATGSCGCTCT CCTTCTGAA GACCCGCTG GAGTGTGGCA TCGGCTTCA 1440
CATCATCTC AGCGGCTGC CCGTCTACTT CTTCGGGCTG TGGTGGAAA ACAAGCCCA 1500
GTGGCTCTC CAGGGCATCT TCTCCAGAC CGTCTGTGT CAGAAGCTCA TGCAGGTGCT 1560
CCCCAGAG ACATAGCCAG GAGGCGAGT GGCTGCCGA GGAGCATGC

Seq ID NO: 87 Protein sequence:
Protein Accession #: XP_035292.2

1 11 21 31 41 51
25 | | | | | |
MAGAGPKRRA LAAPAAEEKE EAREKMLAAK SADGSAPAGE GEGVTIQRNI TLLNGVAIIV 60
GTIIIGSGIFV TPTGVLKEAG SPGLALVWVA ACGVFSIVGA LCYABLGTTI SKSGGDYAYM 120
LEVYGSLEPAF LKLIWELLI RPSSQYIVAF VPATYLLKPL PPTCPVPEEA AKLVACLVL 180
LLTAVNCSYV KAATRVQDAP AAKLLALAL IILGVQVIG KGDVSNLDPN FSFEGTKLDV 240
30 GNIVLALYSY LFAYGWNVYL NFVTEEMINP YRNLEPLAII SLPIVTLVYV LTNLAYFTTL 300
STEQMLSESA VAVDFGNVYL GVMSEIIPVF VGLSCPSGVN GSLETSRLF FVSGREGHLP 360
SILSMIHPLQ LTPVPSLVFT CVMTLLYAFS KDIPSVINFP SFFNWLCLVAL AIIGMTWLRH 420
RKPELEPIK VNLALPVFFI LACLFLIAYS FWKTPVECGI GPTIILSGLP VYFFGVWWRN 480
KPKWLLQIF STTVLCQKLM QVVPQET

Seq ID NO: 88 DNA sequence
Nucleic Acid Accession #: NM_005268.1
Coding sequence: 168-989

1 11 21 31 41 51
40 | | | | | |
TAAAAGCAA AAGAATTCGC GGCCGGGTG ACACGGGCTT CCGGAAAAAC CTTCCTCCGCT 60
TCTGGATATG AAATTCACAG TGCTTGCTGA GTCTATTGCG CGGCTGCTGG GAGCCAGGAG 120
45 AGCCCTGAGG AGTAGTCACT CAGTAGCAGC TGACCGGTGG GTCCACCATT AACTGGAGTA 180
TCTTTGAGG ACTCTGAGT GGGTCAACA AGTACTCCAC AGCCTTTGGG CGCATCTGGC 240
TGTTCTCTGGT CTTCATCTTC CGCGTGCTGG TGTACTCTGT GACGCGCAG CGTGTGTGGA 300
GTGATGACCA CAAGGACTTC GACTGCAATA CTGCGCAGCC CGGCTGCTCT AACGTCTGCT 360
TTGATGAGTT CTTCCTCTGT TCCCATGTGC GCCTCTGGGC CTTGAGCTT ATCCTGGTGA 420
50 CATGCCCTC ACTGCTCGTG GTCATGCAAG TGGCCTACCG GGAGGTTCAG GAGAAGAGGC 480
ACCGAGAAGC CCATGGGAGG AACAGTGGGC GCCTTACCT GAACCCCGGC AAGAAGCGGG 540
GTGGGCTCTG TGGGACATAT GTCTGCAGCC TAGTGTTCAA GGCGAGCGTG GACATGCGCT 600
TTCTCTATGT TTGCACTCA TTCTACCCCA AATATATCCT CCTCTCTGTG GTCAGTGCC 660
AOCGAGATCC ATGTCCCAAT ATAGTGGACT GCTTCATCTC CAAGCCCTCA GAGAAGAAC 720
55 TTTTCACTT CTTCATGAGT GCCACAGCTG CCATCTGCAT CCTGCTCAAC CTGCTGGAGC 780
TCATCTACCT GGTGAGCAAG AGATGCCAGC AGTGCCTGGC AGCAAGGAAA GCTCAAGCCA 840
TGTGACAGG TCATCACCCC CACGGTACCA CCTCTTCTCT CAAACAAGAC GACCTCCTTT 900
CGGGTGAGCT CATCTTCTCT GGCTCAGACA GTCATCTCTC TCTCTTACCA GACCGCCCC 960
GAGACCATGT GAAGAAAACC ATCTTGTGAG GGGCTGCTGG GACTGCTCTG GCAGGTGGG 1020
60 CTTGATGGG GAGGCTCTAG CATCTCTCAT AGGTGCAACC TGAGAGTGGG GGAGCTAAGC 1080
CATGAGGTAG GGGCAGGCAA GAGAGAGGAT TCAGAGCTC TGGAGCCAG TTCTAGTCC 1140
TCAACTCCAG CCACCTGCCC CAGCTCGAGC GCATCGGSC AGTTCCTCCCT CTGCTCTGCA 1200
GCTCGGTTTC CTTTCTAGA ATGGAATAG TGAGGGCCAA TGC

Seq ID NO: 89 Protein sequence:
Protein Accession #: NP_005259.1

1 11 21 31 41 51
70 | | | | | |
MNWSIFEGLL SGVNYSTAF GRIWLSLVFI FRVLVYLVTI ERVWSDDHKD PDCNTRQPGC 60
SNVCFDEFPF VSHVRLWALQ LILVTCPSLL VMHVAYREV QEKHREAHG ENSGRLYLNP 120
GKKRGGWLWT YVCSLVFKAS VDIAPLYVPH SFYPKYILPP VVKCHADPCF NIVDCFISKP 180
SEKNIFTLFM VATAAICILL NLVELIYLVS KRCHCLAAK KAQAMCTGHH PHGTSSCKQ 240
DILLSGDLIF LGSDSHPLL PDRPRDHVKC TIL

Seq ID NO: 90 DNA sequence
Nucleic Acid Accession #: NM_002391.1
Coding sequence: 26-457

1 11 21 31 41 51
80 | | | | | |
CGGGCGAAGC AGCGGGGCA GCGAGATGCA GCACCGAGGC TTCTCTCTCC TCAACCTCTCT 60
CGCCCTGCTG GCGCTCACT CCGGGGTGCG CAAAAAGAAA GATAAGGTGA AGAAGGGCGG 120
85 CCGGGGAGC GAGTGGCTG AGTGGGCTG GGGGCCCTGC ACCCCAGCA GCAAGGATTG 180
CGGGTGGGT TTCCGGGAG GCACCTGGCG GGCCAGAGC CAGCGCATCC GGTGAGGGT 240
GCCCTGCAAC TGGAGAGAG AGTTTGGAG CGACTGCAAG TACAAGTTTG AGAAGTGGG 300
TGCGTGTGAT GGGGACACAG GCACCAAGT CCGCAAGGC ACCCTGAAGA AGGCGCGCTA 360

CAATGCTCAG TGCCAGGAGA CCATCCGGGT CACCAAGCCC TGCACCCCA AGACCAAAAGC 420
 AAAGGCCAAA GCCAAGAAAG GGAAGGGAAA GGAAGGAGAG CCAAGCCTGG ATGCCAAGGA 480
 GCCCTGGTGG TCACATGGGG CCTGGCCACG CCCTCCCTCT CCCAGGCCCG AGATGTGACC 540
 CACCAAGTGC TTCTGTCTGC TCGTTAGCTT TAATCAATCA TGCCTGCTCT TGTCCTCTCT 600
 ACTCCCCAGC CCCACCCCTA AGTGGCCAAA GTGGGGAGGG ACAAGGGATT CTGGGAAGCT 660
 TGAGCTCTCC CCAAGCAAT GTGAGTCCCA GAGCCCGCTT TTGTTCTTCC CCACAATTCC 720
 ATTACTAAGA AACACATCAA ATAAACTGAC TTTTCCCCC CAATAAAAGC TCTTCTTTT 780
 TAATAT

Seq ID NO: 91 Protein sequence:
 Protein Accession #: NP_002382.1

1 11 21 31 41 51
 MQHRGFLLLT LLALLALTSA VAKKKDKVKK GPGSECAEW AWGPCTPSSK DCGVGFREGT 60
 CGAQTQRIKC RVPKNWKEF GADCKYKFEN WGACDGGTGT KVRQGLTKKA RYNAQCQETI 120
 RVTKPCTPKT KAKAKAKGK GKD

Seq ID NO: 92 DNA sequence
 Nucleic Acid Accession #: NM_005130.1
 Coding sequence: 98-802

1 11 21 31 41 51
 CTCTACCTGA CACAGCTGCA GCCTGCAATT CACTCCCACT GCCTGGGATT GCCTGGATC 60
 CGTGTGCTCA GAACAAGGTG AACGCCCAGC TGCAGCCATG AAGATCTGTA GCCTCACCT 120
 GCTCTCCTTC CTCTACTAGG CTGCTCAGGT GCTCCTGGTG GAGGGGAAAA AAAAAAGTAA 180
 GAATGGACTT CACAGCAAAG TGGTCTCAGA ACAAAGGAC ACTCTGGGCA ACACCCAGAT 240
 TAAGCAGAAA AGCAGGCCCG GGAACAAAGG CAAGTTTGTC ACCAAAGACC AAGCCAACTG 300
 CAGATGGGCT GCTACTGAGC AGGAGGAGGG CATCTCTCTC AAGGTTGAGT GCATCAATT 360
 GGACCATGAA TTTTCTGTG TCTTTGCTGG CAATCCAACC TCATGCCTAA AGCTCAAGGA 420
 TGAGAGAGTC TATTGGAAC AAGTTGCCCG GAATCTGCGC TCACAGAAAG ACATCTGTAG 480
 ATATTCCAAG ACAGCTGTGA AAACAGAGT GTGCAGAAAG GATTTTCCAG AATCCAGTCT 540
 TAAGCTAGTC AGCTCCACTC TATTGGGAA CACAAAGCCC AGGAAGGAGA AAACAGAGAT 600
 GTCCCCCAGG GAGCAGATCA AGGCAAAAGA GACCACCCCC TCTAGCCTAG CAGTGACCCA 660
 GACCATGGCC ACCAAAGCTC CCGAGTGTGT GGAGGACCCA GATATGGCAA ACCAGAGGAA 720
 GACTGCCCTG GAGTTCTGTG GAGAGACTTG GAGCTCTCTC TGCACTTCTC TCCTCAGCAT 780
 AGTGACGAGC ACCTCATGCT AATGAGGTCA AAAGAGAACG GGTTCCTTTA AGAGATGTCA 840
 TGTGTAAGT CCCTCTGTAT ACTTTAAAGC TCTCTACAGT CCCCCCAAAA TATGAACITT 900
 TGTGCTAGT GAGTGCAAGC AAATATTTAA ACAAGTTTTC TATTTTGTGC TTTTGTGTTT 960
 TGGAAATTGC CTTATTTTTC TTGGATGCGA TGTTCAGAGG CTGTTTCCTG CAGCATGTAT 1020
 TTCCATGGCC CACACAGCTA TGTGTTGAG CAGCGAAGAG TCTTTGAGCT GAATGAGCCA 1080
 GAGTGATAAT TTCAGTGCAA CGAAGCTTCT GCTGAATTAA TGGTAATAAA ACTCTGGGTG 1140
 TTTTTCAAAA AAAAAAAAAA AAA

Seq ID NO: 93 Protein sequence:
 Protein Accession #: NP_005121.1

1 11 21 31 41 51
 MKICSLTLLS FLLLAQVLL VEGKKKVKNG LHSKVSEQK DTLGNTQIKQ KSRPGNKGKP 60
 VTKDQANCRW AATEQEESIS LKVECTQLDH EFSCVFAGNP TSCLKLKDER VYWKQVARNL 120
 RSQKDICRYS KTAVKTRVCR KDFPESSLKL VSSTLFGNTK PRIKTEMSF REHIKGETT 180
 PSSLAVTQTM ATKAPECVED PDMANQRKTA LEFCGETWSS LCTFFLSIVQ DTSC

Seq ID NO: 94 DNA sequence
 Nucleic Acid Accession #: NM_012101
 Coding sequence: 125-1891

1 11 21 31 41 51
 CTCTCAGCAG GTGTGTCTCT AGTCTCTGTG GTTGCTTGCC CACTCCCTCG CCGAGACGCC 60
 TGCCAGAAAG GTCACCTATC CTGAACCCCA GCAAGCCTGA AACAGCTCAG CCAAGCACCC 120
 TGGATGGAAA GCTGCAGATG CCTCAGGAG CAACGGGTG AGCCAGAGA CCGAGGATGC 180
 CGGAGCCCGC TCGGGCCCCA GTGGCAGCCT GGAGAATGGC ACCAAGGCTG ACGGCAAGGA 240
 TGCCAAGACC ACCAAGCGGC ACGGCGGGGA GGCAGCTGAG GGCAAGAGCC TGGGCAAGGC 300
 CCTGAAGCCA GGGGAAGGTA GGAGCGCCCT GTTCGCGGGC AATGAGTGGC GGGAGCCCAT 360
 CATCCAGTTT GTCGAGTCCG GGGACGACAA GAATCCAAC TACTTCAGCA TGGACTCTAT 420
 GGAAGGCAAG AGGTGCGCGT ACGCAGGGCT CCAGCTGGGG GCTGCCAAGA AGCCACCCGT 480
 TACCTTTGCC GAAAAGGGCG ACGTGCGCAA GTCCATTTTC TCGAGTCCC GGAAGCCAC 540
 GGTGTCCATC ATGGAGCCCG GGGAGACCCG CGGAACAGC TACCCCGGGG CCGACACGGG 600
 CCTTTTTCAT CGGTCCAAGT COGGCTCCGA GGAGGTGCTG TCGACTCTCT GCATCGGCAA 660
 CAAGCAGAAG CGGTCAAGT CCGCTCTGGT GTGCCAGGCC TCCTTCTGCG AGCTGCATCT 720
 CAAGCCCCAC CTGGAGGGCG CCGCTTCCG AGACCCAGC CTGCTCGAGC CCATCGGGT 780
 CTTTGAAGCC CGCAAGTGTG CCGTGATGAG CAAGACGATG GAGCTCTTCT GCCAGACCGA 840
 CCAGACCTGC ATCTGCTACC TTTGATGTT CCAGGAGCAC AAGAATCATA GCACCGTGAC 900
 AGTGGAGGAG GCCAAGGCCG AGAAGGAGAC GGAGCTGTCA CTGCAAAAGG AGCAGCTGCA 960
 GCTCAAGATC ATTGAGATTG AGGATGAAGC TGAGAAGTGG CAGAAGGAGA AGGACCGCAT 1020
 CAAGAGCTTC AACCACTATG AGAAGGCCAT CCGGAGCAG AACTTCGGGG ACCTGGTGG 1080
 GGACCTGGAG AAGCAAAAGG AGGAAGTGAAG GGCTGCGCTG GAGCAGCGGG AGCAGGATGC 1140
 TGTGACCAA GTGAAGGTGA TCATGGATGC TCTGGATGAG AGAGCCAAGG TGCTCATGA 1200
 GGACAGCAG ACCCGGGAGC AGCTGCATAG CATCAGCGAC TGTGTGTTGT TTCTGCAGGA 1260
 ATTTGGTGCA TTGATGAGCA ATTACTTCTC CCCCCACCC CTGCCCACTT ATCATGTCTT 1320
 GCTGGAGGGG GAGGGCCCTG GACAGTCACT AGGCAACTTC AAGGACGACC TGCTCAATGT 1380
 ATGATGCGC CAGCTTGAGA AGATGTGCAA GCGGACCTG AGCCGTAAC TCAATGAGAG 1440
 GAACCAATG GAGAAGCGTG GTGACCATCG CTATGTGAAC AACTACAGA ACAGCTTGG 1500

5
10
15
20
25

```
GGGTGAGTGG AGTGACCCGG ACACCATGAA GAGATACTCC ATGTACCTGA CACCCAAAGG 1560
TGGGGTCCGG ACATCATACC AGCCCTCGTC TCCTGGCGGC TTCACCAAGG AGACCAACCA 1620
GAAGAATTTC AACAAATCTCT ATGGCACCAA AGGTAACATC ACCTCCCGGG TCTGGGAGTA 1680
CTCCTCCAGC ATTCAAGAACT CTGACAATGA CTTGCCCGTC GTCCAAGGCA GCTCCTCCTT 1740
CTCCCTGAAG GGCTATCCCT CCCTCATGCG GAGCCAAAGC CCAAGGGCCC AGCCCCAGAC 1800
TTGGAATCTT GGCAAGCAGA CTATGCTGTC TCACTACCGG CCATTCTACG TGAACAAAGG 1860
CAACGGGATT GGGTCCAACG AAGCCCCATG AGCTCCTGGC GGAAGGAACG AGGCGCCACA 1920
CCCCTGCTCT TCCTCTGAC CTGCTGCTCT TTGCTTCTA AGCTACTGTG CTGTCTCTGG 1980
TGGGAGGGAG CTTGCTCTG CACTGCGCCT CTGACGCGCT CTGCGAGCCT CTGCGGGGCA 2040
GTTCGGGCTT GCTCGACTTC CCCACTGGCC ACACTCCATT CAGACTCCTT TCCTGCTTTG 2100
TGACCTCAGA TGGTCACCAT CATTCCTGTG CTGAGAGGCC AACCCATCAC AGGGGTGAGA 2160
TAGGTTGGGG CTTGCCCTAA CCGGCCAGCC TCCTCCTCTC GGGCTGGATC TGGGGGCTAG 2220
CAGTGAGTAC CCGCATGGTA TCAGCCTGCC TCTCCCGCCC ACGCCTCTGT GTCTCCAGGC 2280
CTATAGACGT TTCTCTCAA GGCCCTATCC CCCAATGTTG TCAGCAGATG CTTGGACAGC 2340
ACAGCCACCC ATCTCCCATT CACATGGCCC ACCTCCTGCT TCCAGAGGA CTGGCCCTAC 2400
GTGCTCTCTC TGTCTTACC TATCAATGCC CAGCATGGCA GAACCTCGAG TGGCCAAAGG 2460
CTGCAGATGG AAACCTCTCA GTGTCTTGAC ATCACCCTAC CCAGGCGGTG GGTCTTCAAC 2520
ACAGCCACTT TGAGTCTGTG GTCCCTGGAG GGTGGCTTCT CTTGACTGGC AGGATGACCT 2580
TAGCCAGATG ATTCTCTGT TCCTCTGCT GAGATAAAGA ATTCCTTAA CATGATATAA 2640
TCCACCATG CAAATGAGT CTGGCCAGC TACCATTTAC CATTTGCCCTA CAGAATTTCA 2700
TTCAGTCTAC ACTTTGGCAT TCTCTCTGCC GATGGAGTGT GGCTGGGCTG ACCGCAAAAG 2760
GTGCTCTACA CACTGCCCCC ACCCTCAGCC GTTGCCCCAT CAGAGGCTGC CTCCTCTTC 2820
TGATTACCCC CCATGTTGCA TATCAGGGTG CTCAAGGATT GGAGAGGAGA CAAAACCAAG 2880
AGCAGCAGAC TGGGGACATC TCCCGTCTCA ACAGCCCGAG GCCTATGGGG GCTCTGGAAG 2940
GATGGGCGAG CTTGACGGGG TTGGGGAGGG AGACATCCAG CTTGGGCTTT CCCCTTTGGA 3000
ATAAACCATT GGTCTGTC
```

Seq ID NO: 95 Protein sequence:
Protein Accession #: NP_036233.1

30
35
40
45

```
1 11 21 31 41 51
MEAADASRSN GSSPEARDAR SPSGPGSLE NGTKADGKDA KTINGHGGEA ABGKSLGSAL 60
KPGEGRSALF AGNEWRRPII QFVESGDDKN SNYFSMDSME GKRSYPAGLQ LGAARKPPVT 120
FAEKGDVRSK IPSESRRKPTV SIMEPGETRR NSYPRADTGL FSRKSGSSEE VLCDSCIGNK 180
QKAVKSLCLV QASFCEHLK PHLEGAAPRD HQLLEPIRDF EARKCPVHKG TMELFCQTDQ 240
TCICVLCMPQ EKHQNSTVTV EBAKAKEETE LSLQKEQLQL KIIEIEDEAE KWQKEKDRIK 300
SPTTNEKAIL EQNPRDLVRD LEKQKEEVRA ALBQREQDAV DQVKVIMDAL DERAKVLHED 360
KQTRQLHSI SDSVLFLOEF GALMSNYSLP PPLPTYHVL EGEGLGQSLG NFKDDLNLVC 420
MRHVEKMCKA DLSRNFIERN HMENGGDHRV VNNYTNSEFG EWSAPDTMKR YSMYLTTPKG 480
VRTSYQPSPP GRFTKETTKQ NFNNLYGTGK NYTSRVWEYS SSIQNSDNDL PVVQSSSFP 540
LKGYPQLMRS QSPKAQPTW KSGKQTMLSH YRPPYVKNKN GIGSNEAP
```

Seq ID NO: 96 DNA sequence
Nucleic Acid Accession #: NM_080668.1
Coding sequence: 83-841

50
55
60
65
70
75
80
85

```
1 11 21 31 41 51
GGCACGAGGG CAGCGAGTGG CTTTCCCGGT TGGCGCGCGC CCGGGCGGCG GCGCTGGAG 60
GAGCTGAGAG CCGAGCCTAG TTATGTCTTG GAGCGGAACG CGGTCCGAG GAGCCGCTCA 120
GCGCTCCGGG CCAAGGGGCC CATCTCCTAC TAAGCCTCTG CCGAGGTCCC AGCGGAATC 180
AGGCTCTGAA CTCGCCAGCA TCCTCCCTGA AATCTGGCCG AAGACACCCA GTGCGGCTGC 240
AGTCAGAAAG CCAATCGTCT TAAAGAGGAT CGTGGCCCAT GCTGTAGAGG TCCCGACTGT 300
CCAATCAGCT CCGAGGAGCC CTAGGATTTC CTTTTTCTTG GAGAAGAAAC ACGAGCCCCC 360
TGGCAGGGAG CTTACTAAGG AGGACCTTTT CAAGACACAC AGCGTCCCTG CCAACCCAC 420
CAGCACTCCT GTGCCGAACC CTGAGGCCGA GTCCAGCTCC AAGGAAGGAG AGCTGGACGC 480
CAGAGACTGT GAAATGTCTA AGAAAGTCAG GCGTTCCTAC AGCCGGCTGG AGACCTGGG 540
CTCTGCTCT CTCTCCACCC CAGGCGCGCG GTCCCTGCTT GGCTTCGAGG GGCTGCTGG 600
GGCAGAGAGC TTGTCCGGAG TCTGCCAGT GGTGTGCTCC AAATCAACG AGGTCCCCAG 660
GGTTTGTGCA AAGCCCTGGG CCCCAGACAT GACTCTCCCT GGAATCTCCC CACCACCGA 720
GAAACAGAAA CGTAAGAAGA AGAAAATGCC AGAGATCTTG AAAACGAGC TGGATGAGTG 780
GGCTCGGCCC ATGAATGCCG AGTTTGAAGC TGCTGAGCAG TTGTATCTCC TGGTTGAATG 840
AGATGCACTG GGGGCTGCAC CTGGCCAGAC TCTCCCTCCT GTCCTGTACA TAGCCACCTC 900
CCTGTGGAGA GGACACTTAG GGTCCCTCC CTTGCTCTTG TTACCTGTGT GTGTGCTGGT 960
GCTGCGCATG AGGACTGTCT GCCTTTGAGG GCTTGGGCAG CAGCGGCAGC CATCTTGGTT 1020
TTAGGAAATG GGGCCGCTTG GCCCAGCCAC TCACTGGTGT CTTGTCTCTT GTCGTCTGT 1080
CCTTCTATC TCCCCAAGT ACCATAGCCA GTTTCAGAT GGGCCACAGA CTGGGGAGGA 1140
GAATCAGTGG CCCAGCCAGA AGTTAAAGGG CTGAGGGTTG AGGTGAGAGG CACCTCTGCT 1200
CTTGTGGGGA GGGGTGGCTG CTTGAAATA GGCCAGGGG CTCTGCCAGC CTCGGCTCT 1260
CCCTCCTGAG TTGCTTCTG TTGTGGCTT TCTTCTTGA CCACTCTGT TAAAGAGGTT 1320
TTCACTTCCG TGGGTTTCCC CTTTGATTCT GTAAATAGTC CCAGAGAGAA TTCGTGGCT 1380
GAGGGCAATT CTGTCTTGA GGAAGAAGCT GGACATTGAG CTTGTGGAGT CTGAGTTTGT 1440
AAGGATGTAG GGGGCTTAG TTGGGTCTCA GACCATAAAT GTGTACTACA CAGAAGCTGT 1500
GTTTCTAGT TCTGCTGTC TGTGAGATG TTTGTAATG GCCAGGTGA TAGGGCGCTG 1560
GCTGCTTGA GCAAAGGGTG CATTTCAGGG TGTGGCCACC AGGTGCTGTG AGTTTCTGTG 1620
GCTCATGGCC TCTGGCTTGG TCCCTTGCAC AGGGCCACG CTGGAGTCTT ACCACTCTGC 1680
TGCAGGGGTG GAAGGTGGCC CCTCTGTCA CCCATACCCA TTTCTTACAA AATAAGTTAC 1740
ACCGAGTCTA CTTGGCCCTA GAGAGAAAG TTGAAGAGTC CCAGACCTAC TAGCATTGTG 1800
CAACTATGCT TGTAAAGTCC TCGGAAAGTT TCCTCGCGTA CCAGACAGCG GCGGGGGCTG 1860
ATAGCAATTT TAGTTTTTGG CCTCCCTATC CTCTCACATG AGAACACTGC CTGGATGCAT 1920
CTCATGATCT CTGGAGAAAT TCCCATCTT TCTTCTCTT CCATGCTGTG GATTCAATAG 1980
TTTGGATTGT AAGGCTGCCC TGCCCCCGAC TCTCTGCGG CACCCCTGGC CATTTGACT 2040
TTTGATGTTT AGAAGTTCGT GGAAAGTAGAC CTGAGGTGT GCAGAGGAGC TGGTGGATAA 2100
CAGAGAATGC CAGGGAAGAT GAGTGTCTGG TCAGGTAATG TGGATGAAC GGTGCAGGCC 2160
AGCGGGGCC TAATAAAACC CTCTGCCAGG TCTGGGAGTC CCAGGCCATC TGCTCAACGC 2220
```


TCTGTGGTTT GTCAGACCTG CAAGCAAGCC CCCTGCTGGG GAAGCCTAGG TGTCTTGTAG 2280
 CTGAACCGCA CTGAAGAACT CTGTCTCTCA CTGGCTGATG CAGCAGAACT CTGGGGAAT 2340
 GTCTTAGTCC TGCAAGATCA GGAGTCACCA GATGATGCAG AGTTGAGATC ATCATTGCAA 2400
 AGTTCTCTGT TCTGTAGGAA CTAATTTTAA GGAAAAAATG GGATTTTGTG TTAGAGTTGG 2460
 AAAAAAAGCC TGATTAAGA GTTTCTGCCT GTTAAAAAAA AAAAAAATAA AAAAAA

Seq ID NO: 97. Protein sequence:
 Protein Accession #: NP_542399.1

1 11 21 31 41 51
 MSGRTRSGG AAQRSGPRAP SPTKPLRSQ RKSGSELPSI LPEIWPKTPS AAARVKPIVL 60
 KRIVAHAVEV PAVQSPRRSP RISFFLEKEN EPPGRELTKS DLFKTHSVPA TPTSTPVPNP 120
 EAESSKKEGE LDARDLEMSK KVRRSYSRLS TLGSASTSTP GRRSCFGFEG LLGAEDLSGV 180
 SPVVCCKLTE PVRVCAKPWA PDMTLPGISP PPEKQKRKKK KMPKILKTEL DEWAAAMNAE 240
 FEAAEQFDLL VE

Seq ID NO: 98 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 58-12444

1 11 21 31 41 51
 GGGGCATTTC CGGGTCCGGG CCGAGCGGGC GCACGCGCGG GAGCGGGACT CGGCGGCATG 60
 GCGGGCTCCG GAGCGGGTGT GCGTTGCTCC CTGCTGCGGC TGCAGGAGAC CTGTGCGCT 120
 GCGGACCGCT GCGGTGCTGC CCTGGCCGGT CATCACTGA TCCGCGGCCCT GGGCAGGAA 180
 TGCTCTCTGA CAGCAGCGCC CGCGTGCTG GCATTACAGA CATCTTAGT TTTTCCAGA 240
 GATTTGCTGT TGCTGTATT TGTCCGGAAG TCACTCAACA GTATTGAAT TCGTGAATGT 300
 AGAGAAGAAA TCTTAAGTT TTTATGTATT TTCTAGAAA AAATGGGCCA GAAGATCGCA 360
 CTTACTCTGT TTGAATTTAA GAACACTTGT ACCAGTGTAT ATACAAAAGA TAGAGCTGCT 420
 AAATGTAAAA TTCCAGCCCT GGACCTTCTT ATTAAGTTAC TTCAGACTTT TAGAAGTTCT 480
 AGACTCATGG ATGAATTTAA AATTGGAGAA TTTATAGTA AATTCTATGG AGAAGTTGCA 540
 TTGAAAAAAA AAATACCGAG TACAGTTTAA GAAAAAGTAT ATGAGCTCTC AGGATTATGT 600
 GGTGAAGTTC ATCTAGTGA GATGATAAAT AATGCAGAAA ACCTGTTCGG CGCTTTCTG 660
 GGTGAACCTTA AGACCCAGAT GACATCAGCA GTAAGAGAGC CCAAACTACC TGTTCGGCA 720
 GGATGTCTGA AGGGGTGTG CTCACTCTG TGCAACTTCA CTAAGTCCAT GGAAGAAAGAT 780
 CCCCAGACTT CAAGGGAGAT TTTTAATTTT GTACTAAAGG CAATTCTGCT TCAGATTGAT 840
 CTGAAGAGAT ATGCTGTGCC CTCAGCTGGC TTGCGCCTAT TTGCCCTGCA TGCATCTCAG 900
 TTTAGCACCT GCCTTCTGGA CAACACGTC TCTCTATTG AGTCTTGT AAAGTGGTGT 960
 GCCACACAAA ATGTAGAATT GAAAAAAGCT GCACCTTCAG CCTTGGAAAT CTTTCTGAAA 1020
 CAGGTTTCTA ATATGGTGGC GAAAAATGCA GAAATGCATA AAAATAAAT GCAGTACTTT 1080
 ATGAGAGCTT TTTATGTTAA CATCAGAAAT GTGGATTGCA ACAACAAGGA GTTATCTATT 1140
 GCTATCCGTC GATATGGACT TTTTGCAGGA CCGTGCAAGG TTAATAACGC AAAAGATGTT 1200
 GACTTCATGT ACCTGTGAGT CATTCAGCGC TGCAAGCAGA TGTCTCTCAC CCAGACAGAC 1260
 ACTGGTGAGC ACCGTGTTTA TCAGATGCCA AGCTTCTCC AGTCTGTGTC AAGCGTCTG 1320
 CTGTACCTTG ACACAGTTCG TGAGGTGTAT ACTCCAGTTC TGGAGCACCT CGTGGTGTG 1380
 CAGATAGACA GTTTCCTTCA GTACAGTCCA AAAATGCAGC TGGTGTGTG CAGAGCCATA 1440
 GTGAAGGTGT TCTAGCTTTT GGCAGCAAAA GGGCCAGTTC TCAGGAATGT CATTAGTACT 1500
 GTGGTGATC AGGGTTTAA CAGAAATATG TCTAAACAG TGGTCTCTCC AAAGGGCCCT 1560
 GAGTCTGAAT CTGAAGACCA CCGTGCTTCA GGGGAAGTCA GAAGTGGCAA ATGGAAGGTG 1620
 CCCACATACA AAGACTAGT GATCTCTTCC AGACATCTCC TGAGCTCTGA CCAGATGATG 1680
 GATTCTATT TAGCAGATGA AGCATTTTTC TCTGTGAATT CCTCCAGTGA AAGTCTGAAT 1740
 CATTTACTTT ATGATGAATT TGTAAATACC GTTTTGAAGA TTGTTGAGAA ATTGGATCTT 1800
 ACATTGAAA TACAGACTGT TGGGAACAA GAGAATGGAG ATGAGGCCCT TGGTGTGTTG 1860
 ATGATCCCAA CTTAGATACC AGCGGCTAAC TTGCATCCAG CTAACCTTAA AGATTTTCTG 1920
 GCTTTTATTA ACCTGGTGGG ATTTTGCGA GAGATTCTCC CTGAGAAACA AGCAGAAATT 1980
 TTTGAACCAT GGGGTGATCT ATTTTCATAT GAATTAATT TGCAATCTAC AAGGTGTGCC 2040
 CTCATCAGTG GTTCTTACAA ATTGCTTTCT ATTACAGTAA GAAATGCCAA GAAATAAAAA 2100
 TATTTGAGG GAGTTAGTCT AAAGAGTCTG AAACACTCTC CTGAAGACCC AGAAAAGTAT 2160
 TCTTGCTTTG CTTTATTGTG GAAATTTGCC AAAGAGGTGG CAGTTAAAT GAAGCAGTAC 2220
 AAAGATGAAC TTTTGGCTCT TTGTTTGACC TTCTTCTGT CCTTGCCACA CAACATCATT 2280
 GAAGTCGATG TTAGAGCTCA CGTTCTCTGA CTGCAGATGG CTTTCAAAT GGGCGCTGAG 2340
 TATACCCCTT TGGCAGTAA AGGCTCTGAT GCTCTAGAAG AATGGTCAAT TTATATTGAC 2400
 AGACATGTAA TGCAGCTTAA TTCAAAAGAC ATTCTCCCTC GCCTGGATGG ATACCTGAAG 2460
 ACTTCAGCCT TGTAGATGA GACCAAGAA AACTGGGAAG TGTAGCTCT TTCTCGGCT 2520
 GCCCAGAAAG GATTTAATAA AGTGGTGTTA AAGCATCTGA AGAAGACAAA GAACCTTCA 2580
 TCAAAACGAG CAATATCCTT AGAAGAAATA AGAATTAGAG TAGTACAAAT GCTTGGATCT 2640
 CTAGGAGGAC AAATAAACAA AAATCTCTG ACAGTCAGT CCTCAGATGA GATGATGAAG 2700
 AGCTATGTG CCTGGGACAG AGAGAAGCGG CTGAGCTTTG CAGTGCCCTT TAGAGAGATG 2760
 AAACCTGTCA TTTCTCTGGA TGTGTTCTCG CCTCGAGTCA CAGAATTAGC GCTCAGACCC 2820
 AGTGACAGAC AAATAAAGT TGCAGCTGT GAACCTTTAC ATAGCATGGT TATGTTTATG 2880
 TTGGGCAAGG CCACGAGAT GCCAGAAGG GGCAGGGGAG CCCCACCAT GTACAGCTC 2940
 TATAAGCGGA CGTTTCTGT GCTGCTTGA CTGGGTGTG ATGTTGATCA GGTGACAGG 3000
 CAAGTGTAT AGCCACTAGT TATGAGCTG ATTCAGTGT TCACTAACAA CAAGAAATTT 3060
 GAAAGTCAGG ATACTGTGTC CTTACTAGAA GCTATATTG ATGGAATGT GAGCCCTGTT 3120
 GACGACTTCT TAAGAGATTT TTGTTGTCGG TGTATTCGAG AATTCCTTAA ATGTTCCATT 3180
 AAGCAATAAA CACCAAGAGA GCAGGAGAAG AGTCCAGTAA ACACCAATC GCTTTTCAAG 3240
 GCATTTTATA GCCTTGGCTG TCACCCCAAT GCTTTCAAGA GGGTGGGAGC ATCACTTGCC 3300
 TTAATAATA TCTCAGGGA ATTCAGGGA GAAGAGTCTC TGGTGAACA GTTTGTGTT 3360
 GAAGCCTTGG TGATATACAT GGAGAGTCTG GCCTTAGCAC ATGCAGATGA GAAGTCTTAA 3420
 GGTACAATCT AACAGTGTG TGATGCCATT GATCACTTAT GCCCATCAT TGAAAAGAG 3480
 CATGTTTCTT TAAATAAAGC AAAGAAACGA CGTTTGGCGG GAGGATTTC ACCTTCGCA 3540
 TCAITGTGTT TATGAGATCT GGTCAAGTGG CTTTATGCTC ATTTGGGGAG GCCCAGACA 3600
 GAAITGTGAC ACAATCCAT TGAACCTTTT TATAAATCG TTCTTTATT GCCAGGCAAC 3660
 AGATCCCTTA ATTTGTGGCT GAAAGATGTT CTAAGGAAG AAGGTGCTCT TTTTCTCATC 3720
 AACACCTTTG AGGGGGGTGG CTGTGGCCAG CCCTGGGGCA TCCTGGGCCA GCCCACCTC 3780
 TTGTACCTTC GGGGGCCATT CAGCCTGCAG GCCAGCTAT GCTGGCTGGA CTGCTCTCTG 3840

	GC0GCGTTGG	AGTGCTACAA	CACGTTTCATT	GGCGAGAGAA	CTGTAGGAGC	GCTCCAGGTC	3900
	CTAGGTACTG	AAGCCAGTC	TTCACCTTTG	AAAGCAGTGG	CTTTCTCTCT	AGAAAGCAIT	3950
	GCCATGCATG	ACATTATAGC	AGCAGAAAAG	TGCTTTGGCA	CTGGGGCAGC	AGGTAACAGA	4020
5	ACAAGCCAC	AAGAGGAGA	AAGGTACAAC	TACAGCAAA	GCACCGTTGT	GGTCCGGATT	4080
	ATGGAGTTTA	CCAGACTCT	GCTAAACACC	TCCCGGAAG	GATGGAAGCT	CCTGAAGAAG	4140
	GACTTGTGTA	ATACACACCT	GATGAGAGTC	CTGGTGCA	CGCTGTGTGA	GCCCGCAAGC	4200
	ATAGGTTTCA	ACATCGGAGA	CGTCCAGGTT	ATGGCTCATC	TTCCTGATGT	TTGTGTGAAT	4260
	CTGATGAAAG	CTCTAAAGAT	GTCCCATAC	AAAGATATCC	TAGAGACCCA	TCTGAGAGAG	4320
10	AAAATAACAG	CACAGAGCAT	TGAGGAGCTT	TGTGCCGTCA	ACTTGTATGG	CCCTGAACGG	4380
	CAAGTGGACA	GGAGCAGGCT	GGCTGCTGTT	GTGTCTGCCT	GTAAACAGCT	TCACAGAGCT	4440
	GGGCTTCTGC	ATAATATATT	ACCGTCTCAG	TCCACAGATT	TGCATCATT	TGTTGGCACA	4500
	GAACCTCTTT	CCCTGGTTTA	TAAAGGCATT	GCOCCTGGAG	ATGAGAGACA	GTGTCTGCCT	4560
	TCTCTAGACC	TCAGTTGTAA	GCAGCTGGCC	AGCGGACTTC	TGGAGTTAGC	CTTTGCTTTT	4620
	GGAGGACTGT	GTGAGCGCTT	TGTGAGTCTT	CTCCTGAACC	CAGCGGTGCT	GTCACCGGG	4680
15	TCCTTGGGCA	GCTCAAGGG	CAGGCTCATC	CACCTCTCCC	ATGGGGAGTA	TTTCTATAGC	4740
	TTGTTCTCAG	AAACGATCAA	CACGGAATTA	TTGAAAATC	TGGATCTTGC	TGTATTGGAG	4800
	CTCATGCAGT	CTTCAGTGA	TAATACCAAA	ATGGTGAAGT	CCGTTTGA	CGGCATGTTA	4860
	GACCAGAGCT	TCAGGGAGCG	AGCAAAACAG	AAACACCAAG	GACTGAAACT	TGCGACTACA	4920
	ATTCCTGCAAC	ACTGGAGAA	GTGTGATTCA	TGGTGGGCCA	AAGATTCCCC	TCTCGAAACT	4980
20	AAAATGGCAG	TGCTGGCCCT	ACTGGCAAAA	ATTTACAGA	TGTATTCACT	TGTATCTTTT	5040
	AATACAGTC	ATGGTTTCATT	CCCTGAAGTC	TTTACAACAT	ATATTAGTCT	ACTTGTCTGAC	5100
	ACAAAGCTGG	ATCTACATT	AAAGGGCCAA	GCTGTCACTC	TTCTTCCATT	CTTCACCAGC	5160
	CTCACTGGAG	GCAGTCTGGA	GGAACTTAGA	CGTGTCTGG	AGCAGCTCAT	CGTTGCTCAC	5220
25	TTCCCATGCG	AGTCCAGGGA	ATTTCTTCCA	GGAACTCCCG	GGTTCAATAA	TTATGTGGAC	5280
	TGCATGAAA	AGTTTCTAGA	TGCATTGGAA	TTATCTCAA	GCCTATGTT	GTTGGAATTG	5340
	ATGACAGAAG	TTCTTTGTGG	GGAAACAGCAG	CATGTCTAGG	AAGAAATTAT	TCAATCCAGT	5400
	TTCAAGGAGA	TTGCCAGAG	GGGTTTCATGT	GTACACAAAG	TAGGCTTCT	GGAAAGCGTG	5460
	TATGAATGT	TCAGGAAGGA	TGACCCCGCG	CTAAGTTTCA	CACGCCAGTC	CTTTGTGGAC	5520
	CGCTCCCTCC	TCATCTGCT	GTGGCACTGT	AGCCTGGATG	CTTTGAGAGA	ATTCCTTCAGC	5580
30	ACAATTGTGG	TGGATGCCCT	TGATGTGTTG	AAGTCCAGGT	TTACAAAGCT	AAATGAATCT	5640
	ACCTTTGATA	CTCAAACTAC	CAAGAAGATG	GGCTACTATA	AGATTCTAGA	CGTGTATGAT	5700
	TCTCGCCTTC	CCAAAGATGA	TGTTTCATGCT	AAGGAATCAA	AAATTAACTA	AGTTTTCAT	5760
	GGCTCGTGTA	TTACAGAAGG	AAATGAACCT	ACAAAGACAT	TGATTTAAAT	GTGCTACGAT	5820
	GCATTACAG	AGAATCATGG	AGGAGAGAAT	CAGCTGTGCG	AGAGGAGAAG	ACTTTACCAT	5880
35	TGTGCAGCAT	ACAATCGGC	CATATCTGTC	ATCTGCTGTG	TCTTCAATGA	GTTAAATTT	5940
	TACCAAGGTT	TTCTGTTTAG	TGAAAAACCA	GAAGAAGACT	TGCTTATTTT	TGAAATCTG	6000
	ATCGACCTGA	AGCGCCGCTA	TAATTTTCCT	GTAGAAGTTG	AGGTTCTCTAT	GGAAAGAAAG	6060
	AAAAAGTACA	TTGAAATTAG	GAAAGAACCC	AGAGAAGCAG	CAAAATGGGA	TTCAAGTGGT	6120
	CCTTCCCTATA	TGCTTCCCT	GTCAATTTG	GCAGACAGTA	CCCTGAGTGA	GGAAATGAGT	6180
40	CAATTTGATT	TCTCAACCG	AGTTCAAGAG	TATTCATACA	GCTCCCAAGA	CCCTAGACCT	6240
	GCCACTGGTC	TTTTCGGAG	ACGGGAGCAG	CGGAGCCCA	CGGTGCATGA	TGATGTGCTG	6300
	GAGCTGGAGA	TGGACGAGCT	CAATCGGCAT	GAGTGCATGG	CGCCCTGAC	GGCCCTGGTC	6360
	AAGCACAATG	ACAGAAGCTC	GGGCCCCCT	CAAGGAGAAG	AGGATTCACT	GCCCAAGAGAT	6420
	CTTCTCTCT	GGATGAAAT	CCTCCATGGC	AACTGGGAA	ATCCAATAGT	ACCAATTAAT	6480
45	ATCCGCTCT	TCTAGCCAA	GCTTGTATT	AATACAGAAG	AGGCTTTTGG	CCCTTACGGG	6540
	AAGCACTGGC	TTAGCCCTT	GCTGCAGCTG	GCTGCTCTG	AAAACAATGG	AGGAGAAGGA	6600
	ATTCACTACA	TGGTGGTTGA	GATAGTGGCC	ACTATTCTTT	CATGGACAGG	CTTGCCCACT	6660
	CCAAAGGGG	TGCTTAAAGA	TGAAGTGTTA	GCAAACTGAT	TGCTTAAATT	CCTAATGAAA	6720
	CATGTCTTCT	ATCCAAAAG	AGCTGTGTTT	AGACACAACC	TTGAAATTAT	AAAGACCTTT	6780
50	TGCGAGTGCT	GGAGGATG	TTTATCCATC	CCTTATAGT	TAATATTGGA	AAAGTTTTC	6840
	GGTAAAGATC	CTAATTCTAA	AGACAACCTCA	GTAGGGATTC	AATTGCTAGG	CATCGTGTG	6900
	GCCAAAGACC	TGCTTCCCTA	TGACCCACAG	TGTGGCATCC	AGAGTAGGCA	ATACTTCCAG	6960
	GCTTTGGTGA	ATAATATGTC	CTTTGTAAGA	TATAAAGAA	TGTATGCCCG	TGCAGCAGAA	7020
	GTTCTAGGAC	TTATACTTGG	ATATGTTATG	GAGAGAAAA	ACATCTGGA	GGAGTCTCTG	7080
55	TGTGAACCTG	TGCGAAACA	ATTGAAGCAA	CATCAGAATA	CTATGGAGGA	CAAGTTTATT	7140
	GTGTGCTTGA	ACAAAGTGAC	CAAGAGCTTC	CCTCCTCTG	CAGACAGGTT	CATGAATGCT	7200
	GTGTCTTTTC	TGCTGCCAAA	ATTTTCATGGA	GTGTGAAAA	CACCTCTCT	GGAGGTGTTA	7260
	CTTTGCTGTC	TGGAGGGAAT	GACAGAGCTG	TACTTCCAGT	TAAAGAGCAA	GGACTCTCGT	7320
	CAAGTCACTA	GACATAGAGA	TGATGAAAGA	CAAAAGTAT	GTTTGGACAT	AATTTATAAG	7380
60	ATGATGCCAA	AGTTTAAACC	AGTAGAACTC	CGAGAACTTC	TGAACCCCTG	TGTGGAATTC	7440
	GTTTCCCATC	CTTCTACAAC	ATGTAGGGAA	CAAAATGTATA	ATATTCTCAT	GTGGATTTCAT	7500
	GATAATTACA	GAGATCCAGA	AAGTGAGACA	GATAATGACT	CCCAGGAAAT	ATTTAAGTTG	7560
	GCAAAAGATG	TGCTGATTCA	AGGATTGATC	GATGAGAACC	CTGGACTTCA	ATTAATTATT	7620
	CGAAATTTCT	GGAGCCATGA	AACATAGTTA	CCTTCAAATA	CCTTGGACCG	GTTGCTGGCA	7680
65	CTAAATTCCT	TATATTCTCC	TAAAGTAGAA	GTGCACTTTT	TAAGTTTAGC	AACAAATTTT	7740
	CTGCTCGAAA	TGACCAGCAT	GAGCCAGAT	TATCCAAACC	CCATGTTTGA	GCATCTCTGT	7800
	TCAGAAATGCG	AATTTACAGGA	ATATACCATT	GATTCTGATT	GGCGTTTCCG	AAGTACTGTT	7860
	CTCACTCOGA	TGTTTGTGGA	GACCCAGGCC	TCCAGGGCA	CTCTCCAGAC	CCGTACCCAG	7920
	GAAGGGTCCC	TCTCAGCTCG	CTGGCCAGTG	GCAGGGCAGA	TAAGGGCCAC	CCAGCAGCAG	7980
70	CATGACTTCA	CATGACACA	GACTGCAGAT	GGAAGAAGCT	CATTGATTG	GCTGACCGGG	8040
	AGCAGCACTG	ACCGCTGTGT	CGACCACACC	AGTCCCTCAT	CTGACTCCTT	GCTGTTTGCC	8100
	CACAGAGGGA	GTGAAAGGTT	ACAGAGAGCA	CCCTTGAAGT	CAGTGGGGCC	TGATTTTGGG	8160
	AAAAAAGGCC	TGGGCTTCC	AGGGGACGAG	GTGGATAACA	AAGTGAAAGG	TGCGGCCCGC	8220
	CGGACGGACC	TACTACGACT	CGGCAGACGG	TTTATGAGGG	ACCGAGGAGAA	GCTCAGTTTG	8280
75	ATGTATGCCA	GAAAGGGT	TGCTGAGCAA	AAACGAGAGA	AGGAAATCAA	GAGTGAGTTA	8340
	AAAAATGAAG	AGGATGCCCA	GGTCTTCTG	TACAGAAGCT	ACCGGCAACG	AGACCTTCTC	8400
	GACATTGAGA	TCAAGCAGAG	CAGCTCATC	ACCCGTTTAC	AGGCGGTGGC	CCAGAGGGAC	8460
	CCAATAATTG	CAAAACAGCT	CTTTAGCAGC	TTGTTTCTG	GAATTTTGA	AGAGATGGAT	8520
	AAATTTAAGA	CACGTCTGA	AAAAACAAC	ATCACTCAA	AGTTGCTTCA	AGACTTCAAT	8580
80	CGTTTTCTTA	ATACCACCTT	CTCTTCTTT	CCACCTTTG	TCTCTGTAT	TCAGGACATT	8640
	AGCTGTGAGC	AGCGAGCCCT	GCTGAGCCTC	GACCCAGCGG	CTGTTAGCGC	TGGTTGCTG	8700
	GCCAGCCTAC	AGCAGCCCTG	GGGCATCCGC	CTGCTAGAGG	AGGCTCTGCT	CCGCTGCTG	8760
	CCTGCTGAGC	TGCTGTCCAA	GGGAGTCCGT	GGGAAGGCC	GCCTCCCTCC	TGATGTCTTC	8820
	AGATGGGTGG	AGCTTGCTAA	GCTGTATAGA	TCAATTGGAG	AATACGAGCT	CCTCCGTGGG	8880
85	ATTTTTACCA	GTGAGATAGG	AACAAAGCAA	ATCACTCAGA	GTGCATTATT	AGCAGAAGCC	8940
	AGAAGTGATT	ATTTCTGAAG	TGCTAAGCAG	TATGATGAGG	CTCTCAATAA	ACAAGACTGG	9000
	GTAGATGGTG	AGCCCCACAG	AGCCGAGAAG	GATTTTGGG	AACCTGCATC	CCTTGACTGT	9060

TACAACCAACC TTGCTGAGTG GAAATCACTT GAATACTGTT CTACAGCCAG TATAGACAGT 9120
 GAGAACCCCC CAGACCTAAA TAAATCTGG AGTGAAACAT TTTATCAGGA AACATATCTA 9180
 CCTTACATGA TCCGCAGCAA GCTGAAGCTG CTGCTCCAGG GAGAGGCTGA CCAGTCCCTG 9240
 CTGACATTTA TTGACAAAGC TATGCACGGG GAGCTCCAGA AGGGGATTCT AGAGCTTCAT 9300
 TACAGTCAAG AGCTGAGTCT GCTTTACCTC CTGCAAGATG ATGTTGACAG AGCCAAATAT 9360
 TACATTCAAA ATGGCATTCA GAGTTTATG CAGAATTATT CTAGTATTGA TGTCTCTTAA 9420
 CACCAAGAAT GACTCACCAAT ATTGCAGTCT GTACAGGCTT TAACAGAAAT TCAGGAGTTC 9480
 ATCAGCTTTA TAAGCAACA AGGCAATTTA TCATCTCAAG TTCCCTTTAA GAGACTTCTG 9540
 AACACCTGGA CAACAGATA TCCAGATGCT AAAATGGACC CAATGAACAT CTGGGATGAC 9600
 ATCATCACAA ATCGATGTTT CTTTCTCAGC AAAATAGAGG AGAAGCTTAC CCCTCTTCCA 9660
 GAAGATAATA GTATGAATGT GGATCAAGAT GGAGACCCCA GTGACAGGAT GGAAGTGCAA 9720
 GAGCAGGAAG AAGATATCAG TCCTCTGATC AGGAGTTGCA AGTTTCCAT GAAAATGAAG 9780
 ATGATAGACA GTGCCCGGAA GCAGAACAA TTTCTCACTG CTATGAAACT ACTGAAGGAG 9840
 CTGCATAAAG AGTCAAAAC CAGAGACGAT TGGCTGGTGA GCTGGGTGCA GAGCTACTGC 9900
 CGCTGAGCC AGCTCGGAG CCGGTCCAG GCGTCTCTG AGCAGGTGCT CACTGTGCTG 9960
 AAAACAGTCT CTTTGTGGA TGAGAACAA CTGTCAAGCT ACTTAAGCAA AAATATTCTG 10020
 GCCTTCGGT ACCAGAACAT TCTCTGGGT ACAACTTACA GGATCATAGC GAATGCTCTC 10080
 AGCAGTGAGC CAGCTCCCTT TGCTGAAATC GAGGAGGACA AGGCTAGAAG AATCTTAGAG 10140
 CTTTCTGGAT CCAATTCAGA GGATTCAGAG AAGGTGATCG CGGGTCTGTA CCAGAGAGCA 10200
 TTCCAGCACC TCTCAGGCG GTGCAAGGCG GCTGAGGAGG AGGCCAGCC TCCCTCTCTG 10260
 AGCTGTGGCG CTGCAGCTGG GGTGATTGAT GCTTACATGA CGCTGGCAGA TTTCTGTGAC 10320
 CAACAGCTCG GCAAGGAGGA AGAGAATGCA TCAGTTATTG ATTTCTGAGA ACTGCAGCGC 10380
 TATCCAGCAC TTGTGTGGA GAAAATGTTG AAAGCTTTAA AATTAAATTC CAATGAAGCC 10440
 AGATTGAAGT TTCTAGATT ACTTCAGATT ATAGAACGGT ATCCAGAGGA GACTTTGAGC 10500
 CTGATGACAA AAGAGATCTC TTCCGTTCCC TGCTGGCAGT TCATCAGCTG GATCAGCCAC 10560
 ATGGTGGCT TACTGGACAA AGACCAAGCC GTTGTCTGTC AGCACTCTGT GGAAGAAATC 10620
 ACTGATAACT ACCCGCAGC TATTGTTTAT CCCTTCATCA TAAGCAGCGA AAGCTATTCC 10680
 TTCAAGGATA CTCTACTGG TCATAAGAAT AAGGAGTTTG TGGCAAGGAT TAAAGTAAG 10740
 TTGGATCAAG GAGGAGTGAT TCAAGATTTT ATTAATGCGT TAGATCAGCT CTCTAATCTC 10800
 GAACTGCTCT TTAAGAGTTG GAGCAATGAT TAAAGAGCTG AACTAGACAA AACCCCTGTA 10860
 AATAAAACAA ACATTGAAA AATGTATGAA AGAATGTATG CAGCCTTGGG TGACCCAAAG 10920
 GCTCCAGGCC TGGGGGCTTT TAGAAGGAAG TTTATTGAGA CTTTGGAAA AGAATTGGAT 10980
 AAACATTTTG GGAAGGAGG TTCTAAACTA CTGAGAATGA AGCTCAGTGA CTTCAACGAC 11040
 ATTAACACAA TGCTACTTTT AAAAATGAAC AAAGACTCAA AGCCCCCTGG GAATCTGAAA 11100
 GAATGTTTAC CCGGATGAG CGACTTCAA GTGGAGTTCC TGAGAAATGA GCTGGAGATT 11160
 CCGGTGAGT ATGACGGTAG GGGAAAGCCA TTGCCAGAGT ACCACGTGCG AATCGCCGGG 11220
 TTTGATGAGC GGGTGACAGT CATGGCTCTC CTGCGAAGGC CCAAGCGCAT CATCATCGT 11280
 GGCATGAGC AGAGGGAACA CCCTTCTCTG GTGAAGGGTG GCGAGGACCT GCGGAGGAC 11340
 CAGCGCGTGG AGCAGCTCTT CCAGGTCATG AATGGGATCC TGGCCCAAGA CTCCGCTGCG 11400
 AGCCAGAGGG CCTCTCAGCT GAGGACCTAT AGCGTTGTGC CCATGACCTC CAGGTTAGGA 11460
 TTAATGTAGT GGGTGAAGA TACTGTTACC TTGAAGGACC TTCTTTTGAA CACCATGTCC 11520
 CAAGAGGAGA AGCGGCTTTA CTTGAGTGAT CCGAGGACAC CGCGGTGTA ATATAAGAT 11580
 TGGCTGACAA AAATGTGAGG AAAACATGAT GTTGGAGCTT ACATGCTAAT GTATAAGGGC 11640
 GCTAATCGTA CTGAACAGCT CAGTCTTTT AGAAAACGAG AAAGTAAAGT GCCTGCTGAT 11700
 CTCTTAAAGC GGGCTTCTGT GAGGATGATG ACAGCCCTG AGGCTTTCTT GCGCTCCGCG 11760
 TCCACTTGG CCACTCTCTA CGCTCTGATA TGATCAGCC ACTGGATCCT CGGGATTGGA 11820
 GACAGACATC TGAACAACAT TATGGTGGCC ATGGAGACTG CGCGGTGAT CGGGATCGAC 11880
 TTTGGGCATG CGTTGGATC CGCTACACAG TTTCTGCCAG TCCTGAGTGT GATGCTCTTT 11940
 CGGTAACCTC GCGCACTTAT CAATCTGATG TTACCAATGA AAGAAACGGG CCTTATGTAC 12000
 AGCATCATGG TACACGCTC CCGGGCCTTC CGCTCAGACC CTGGCTGCTC CACCAACACC 12060
 ATGGATGTGT TTGTCAAGGA GCGCTCCTTT GATTGGAAAA ATTTGAACA GAAAAATGCTG 12120
 AAAAAAGAGG GTTCATGGAT TCAAGAAATA AATGTTGCTG AAAAAAATTG GTACCCCGGA 12180
 CAGAAATAT GTTACGCTAA GAGAAAGTTA GCAGGTGCCA ATCCAGCAGT CATTACTTGT 12240
 GATGAGTAC TCCTGGGTCA TGAGAAGGCC CCGCTCTTCA GAGACTATGT GGTGTGGCA 12300
 CGAGGAAGCA AAGATCAACA CATTCTGACC CAAGAACCAG AGAGTGGGCT TCGAAGAG 12360
 ACTCAAGTGA AGTGCTGAT GAGACAGGCA ACAGACCCCA ACATCCTTGG CAGAACCTGG 12420
 GAAGGATGG AGCCCTGAGT GTGAGGTCTG TGGAGTCTG CAGATAGAAA GCATTACATT 12480
 GTTAAAGAA TCTACTATAC TTTGGTTGGC AGCATTTCCAT GAGCTGATTT TCCTGAAACA 12540
 CTAAGAGAAA ATGCTCTTTG TGCTACAGTT TCGTAGCATG AGTTTAAATC AAGATTATGA 12600
 TGAGTAAATG TGATGGGTTT AAATCAAAGA TAAGTTTATA GTAAATCAAA AGATTAGGTG 12660
 AGGTTTATAG AAAGATAGAT ATCCAGGCTT ACCAAAGTAT TAAGTCAAGA ATATAATATG 12720
 TGATCAGCT TCAAGCATCT TACAAGTGCT GCAAGTTAGT GAAACAGCTG TCTCCGTAAA 12780
 TGGAGGAAAT GTGGGAAGC CTTGGAATGC CTTCTGGTT CTGGCATT GGAAGCACA 12840
 CTCAGAGGC TTATCACCA AGATTTTGGG AGAGTAAAGC TAAGTATAGT TGATGTAACA 12900
 TTGTAGAAGC AGCATAGGAA CAATAAGAAC AATAGTAAA GCTATAATTA TGGCTTATAT 12960
 TTAGAAATGA CTGCTTTTGA TATTTTAGGA TATTTTCTA GGTTTTTC TTTTATTATA 13020
 TTTCTTCTA GTTTGACAT TTTATGATAG ATTTGCTCTC TAGAAGGAAA CGTCTTTATT 13080
 TAGGAGGGA AAAATTTTGG TCATAGCATT CACTTTTGTCT ATTCCAATCT ACACTGGAA 13140
 GATACATAAA AGTGCTTGGC ATTTGAATTTG GGATAACTTC AAAAATCCCA TGGTTGTTGT 13200
 TAGGGATAGT ACTAAGCATT TCAGTTCCAG GAGAATAAAA GAAATTCCTA TTTGAAATGA 13260
 ATTCCTCATT TGGAGGAAA AAAGCATGCA TTCTAGCACA ACAAGATGAA ATTATGGAAT 13320
 ACAAAAGTGG CTCTTCCCA TGTGAGTCC CTGTCCCCCC CCGCAGTCC TCACACCCA 13380
 AACTGTTTCT GATTGGCTTT TAGCTTTTGT TGTGTTTTTT TTTCTCTCT AACACTTGTA 13440
 TTTGGAGGCT CTCTGTGAT TTTGAGAAGT ATACTCTTGA GTGTTTAATA AAGTTTTTTT 13500
 CCAAAAGTA

Seq ID NO: 99 Protein sequence:
 Protein Accession #: NP_008835.5

80 1 11 21 31 41 51
 MAGSGAGVRC SLRLQETLS AADROGAALA GHQIRLGLGQ ECVLSSSPAV LALQTSLVPS 60
 RDPGLLVFVR KSLNSIEPRE CREELKFLC IFLEKMGQKI APYSVEIKNT CTSVYTKDRA 120
 ARCKIPALDL LKILQTFRS SRLMDEPKIG ELFSKFYGL ALKKKIPDVT LEKVYELLGL 180
 85 LGEVHPSEMI NNAENLPRAF LGELKTQMTS AVREPKLPVL AGCLKGLSSL CNFTKSMEE 240
 DPQTSREIFN FVLKAIRPOI DLKRYAVPSA GLRLPALHAS QFSTCLLDNY VSLFVLLKW 300
 CARTNVELKK AALSALSFLL KQVSNMVAKN AEMHKNKLQY FMEQPYGIIR NVDSNNKELS 360

5 IAIRGYGLFA GPCKVINAKD VDFMYVELIQ RCKQMPLTQT DTGDDRVYQM PSFLQSVASV 420
 LLYLDTPVEV YTPVLEHLVV MQIDSPPQYS PKMQLVCCRA IVKVFLALAA KGPVLRNCIS 480
 TVVHQGLIRI CSKPVVLPKG PESESEDHRA SGEVRTGKWK VPTYKDYVDL FPHLLSSDQM 540
 MDSILADEAF FSVNSSSESL NHLLYDEFVK SVLKIVEKLD LTLBIQTVEG QENGDEAPGV 600
 WMIPTSDPAA NLHAPAKPKDF SAFINLVEFC REILPEKQAE FFEPMVVSFS YELILQSTRL 660
 FLISGFYKLL SITVRNAKKI KYFEGVSPKS LKESPEDPER YSCFALFVKF GKEVAVRMKQ 720
 YKDELLASCL TPLLSLPHNI IELDVRAIYP ALQMAFKLGL SYTPLAESVGL NALEEWSIYI 780
 DRHVMQPYRK DILPCLDGYL KTSALSDETR NNWEVSALSR AAQKGFNKVV LKHLKTKNL 840
 SSNEAISLEE IRIRVVQMLG SLGGQINKNL LTVTSSDEMM KSYVANDREK RLSFAVPFRE 900
 10 MKPFILODFV LPRVTEALAT ASDRQTKVAA CELLSMVMF MLGKATOMPE GGOGAPPMYQ 960
 LYKRTFPVLL RLACDQDVQT RQLYEPLVMQ LIHWFTNNKK FESQDTVALL EAILDGIQVDP 1020
 VOSTLRDFCG RCIREPLKWS IKQITPQQE KSPVNTKSLF KRLYSLALHP NAFKRLGASL 1080
 AFNNIYREFR EEESLVEQVY FEALVIYMS LALAHADKXS LGTIQQCCDA IDHLCRIIEK 1140
 KHVSLNKKAK RRLPRGFPPS ASLCLLDLVK WLLAHCGRPQ TECHRSIEL FYKFPVLLPG 1200
 15 NRSFNLWLDK VLKKEGVSLF INTFEGGGCG QPSGILAQPT LLYLRGPFSL QATLCWLDLL 1260
 LAALCEYNTF IGERTVAGALQ VLGTEAQSSL LKAVAFFLES IAHDI IAAE KCPGTGAAGN 1320
 RTSPPQGEYR NYSKCTVVVR IMEFTTLLN TSPEGWKLLK KDLCNTHLMR VLVQTLCEPA 1380
 SIGFTIDLVK VMARLPDVC NLKMKALKMSP KYDILETHLR EKITAQISIE LCAVNLVGPD 1440
 AQVDRSLAA VVSACKQLHR AGLLHNILPS QSTDLLHVSQ TELLSLVYKG IAPGDERQCL 1500
 20 PSLDLSCKQL ASGLLELAF PGGLCERLVS LLLNPAVLST ASLGSQGSV IHFSHGEYFY 1560
 SLFSETINTE LLKNLDLAVL ELMQSSVDNT KMVSAVLNGM LDQSFREAN QKHQGLKAT 1620
 TILQHWKCKD SWAKDSPLK TKMAVLALLA KILQIDSSVS FNTSHGSPFE VFTTYISLLA 1680
 DTKLDLHLGK QAVTLLEPFT SLTGGSLLEL RRVLEQLIVA HPMQSRFPF PGTFRFNNV 1740
 DCMKFLDAL RLSQSPMLLE LMTTEVLCREQ QHVEELFQS SFRRIRARRS CVTVQGLLES 1800
 25 VYEMFRKDDP RLSPTQSFV DRSLTLLEWH CSLDALREFF STIUVDAIDV LKSRFTKLE 1860
 STFDQTITKK MGYKILDMV YSRLPKDDVH AKESKINQVF HGSCITEGNE LTKTLIKLCY 1920
 DAFTEMNAGE NQLERRRLR HCAAYNCAIS VICCVFNLK PYQGLFSEK PEKNLLIFEN 1980
 LIDLKRRYNF FVEVEVEMER KKKYIEIRKE AREAANGDSV GPSYMSLSY LADSTLSEEM 2040
 SQDFSTGQV SYSYSQDPR PATGRFRRE QRDPTVHDDV LELEMDLNR HECMAPLTAL 2100
 30 VKHMRSLGF PQGEEDSVPR DLPSWMKFLH GKLGNPVPL NIRFLAKLV INTEEVFRPY 2160
 AKHWSPLQ LAASENNGE GIHYMVVEIV ATILSWTGLA TPTGVPKDEV LANRLNPLM 2220
 KHVFPKRAV PRNLEIITK LVCEWKDCLS IPYRLIFKEP SGKDPNSKDN SVGIQLLGIV 2280
 MANDLPYDP QVQSSEYF QALVNNMSFV RYKEVYAAA EVLGLILRV MERKNILES 2340
 35 LCELVAQKQK QHNTMEDKF IVCLNKVTKS FPPLADRFM AVFFLEPKFH GVLTLCLEV 2400
 VLCRVGME TE LYFLKSKDF VQVMRHRDDE RQKVCIDIY KMKPKLPVE LRELLNPFVE 2460
 FVSHPTTTCR EQMYNIMWI HDNYRDPSE TDNDSEIPEK LAKDVLIQGL IDENPGLQLI 2520
 TRNFWSHETR LPSNTLDRLL ALNSLYSPKI EVHFLSLATN FLEMTSMSP DYPNPMFEHP 2580
 LSECEQEYF IDSDFRFRST VLTMPFVETQ ASQGTQLQTR QEGSLARWP VAGQIRATQ 2640
 40 QHDTTLTQTA DGRSSFDWLT GSSTDPLVDH TSPSSDRLF AHKRSERLQ APLKSVGPDF 2700
 GKRLGLPGD EVDNKVKGAA GRTDLLRLR RFMRDQEKLS LMYARKGVAB QKREKEIKSE 2760
 LKMKQDAQV LYRSYRHGD PDIQIKHSSL ITPQAVAR DPIIAKQFVS SLFSGILKEM 2820
 DKFKTLSEEN NITQKLLQDF NRFLNTTFSF FPPFVSCIQD ISCQHAALLS LDPAAVSAGC 2880
 LASLQQPVG RLLEEALLRL LPAELPAKRV RGKARLPDVP LRWVELAKLY RSIGYDVLR 2940
 45 GIFTSEIGTK QITQSALLAE ARSDYSEAK QYDEALNKQD WVDGEPTAE KDFWELASLD 3000
 CYNBLAEWS LEYCSTASID SENPFDLNI WSEPFYQETY LPMYIRSKLK LLLQGEADQS 3060
 LLTFIDKAMH GELOKALIEL HYSQELSLLY LLQDDVDRAK YIIQNGIQSF MKNYSSIDVL 3120
 LHQSRLTKLT SVQALTEIQE FIFISKQGN LSSQVPLKRL INTWTRYPD AKMDPMNIWD 3180
 DIITNRCFFL SKIEELTLPL PEDNSMNVQD DGDPSDRMEV QEQEEDISSL IRSCKFSMKM 3240
 50 KMDSARKON NFSLAMKLLK ELHKESKTRD DWLVSWVQSY CRLSHCRSRS QGCSEQVLTV 3300
 LKTVSLDEN NVSSYLSKNI LAFRDQNIL GTTYRIIANA LSSEAPCLAB IEEDKARRIL 3360
 ELSSGSSSED EKVYAGLYQR AFQHLSEAVQ AAESEAPPS WSCGPAAGVI DAYMTLADFC 3420
 DQQLRKEEN ASVIDSAEL AYPALVVERM LKALKNSNE ARKCFPRLLQ IIERYPEETL 3480
 SLMTKEISSV PCVPTSLWIS HMAVLLDKQD AVAVQHSVEE ITDNYPQAVI YPFIISSESY 3540
 55 SFKDTSTGKX NKEFVARIKS KLDQGGVIOQ FINALDQSN PELLFKDWSN DVRAELAKTP 3600
 VNKKIEKNY ERYAALGDP KAPGLGAPRR KPIQTGKEF DKHFGGKGL LRLMKLSDFN 3660
 DITNMLLKM NKDSKPPGNL KECSPWMSDF KVEPLRNELE IPGQYDGRGK PLPEYHVRIA 3720
 GFDERVTMA SLRRPKRIII RGHDEREHPF LVKGGEDLRQ DQVEQLFQV MNGILAQDSA 3780
 CSQRLQKLT YSVVPMTRSL GLIEWLENTV TLKDLLNTM SQEKAAYLS DPRAPPCEYK 3840
 60 DWLTKMSGH DVGAYMLMYK GANRTETVTS FRKRESKIPA DLLKRAFVRM STSPEAFAL 3900
 RSHFASHAL ICTISHLGI GDRHLNFMV AMETGGVIGI DFGHAFGSAT QLEVPPELMP 3960
 FRLTRQFINL MLPMKETGLM YSIMVHALRA FRSDPGLLTN TMDVFKPEPS FDKNFEQKM 4020
 LKKGGSWQIE INVAENWYYP RQKICYAKRK LAGANPAVIT CDELLGHEK APAFRDYVAV 4080
 ARGSKDHRIR AQEPESGLSE ETQVKCLMDQ ATDPNILGRT WEGWEPWM

65 Seq ID NO: 100 DNA sequence
 Nucleic Acid Accession #: NM_000673
 Coding sequence: 101-1225

70 1 11 21 31 41 51
 ATGTGAAGGC ACAAGCTGCT GTTATATACA ACAGAGTGAA CTGAGCATCA GTCAGAAAAA 60
 GTCTATGTTT GCAGAAATAC AGATCCAAGA CAAGACAGG ATGGGCACTG CTGAAAAAGT 120
 TATTAATGTC AAAGCAGCTG TGCTTTGGGA GCAGAAGCAA CCTTCTCCA TTGAGGAAAT 180
 75 AGAAGTGGC CCACCAAGA CTAAAGAAGT TCGCATTAAG ATTTTGCCCA CAGGAATCTG 240
 TCGCACAGAT GACCATGTGA TAAAGGAAC AATGGTGTC AAGTTCCAG TGATTGTGGG 300
 ACATGAGGCA ACTGGGATG TAGAGAGCAT TGGAGAAGGA GTGACTACAG TGAACCAGG 360
 TGACAAAGTC ATCCCTCTCT TTCTGCCACA ATGTAGAGAA TGCAATGCTT GTGCAACCCC 420
 AGATGGCAAC CTTTGCAITTA GGAGCGATAT TACTGGTCTG GGAGTACTGG CTGATGGCAC 480
 CACCAGATTT ACATGCAAGG GCAAACCACT ACACCACTTC ATGAACACCA GTACATTATC 540
 80 CGAGTACACA GTGGTGGAT AATCTTCTGT TGCTAAGATT GATGATGCAG CTCTCTCTGA 600
 GAAAGTGTGT TTAATTGGCT GTGGGTTTTC CACTGGATAT GGGCTGCTGT TTAACCTCG 660
 CAAGGTCAAA CCTGGTTCCA CTTCGCTGCT CTTTGGCCTG GGAGGAGTTG GCCTGTCACT 720
 CATCATGGGC TGTAAAGTCAG CTGGTGATC TAGGATCATT GGAGTATGAC TCAACAAAGA 780
 CAAATTGAG AAGGCCATGG CTGTAGGTGC CACTGAGTGT ATCAGTCCCA AGGACTCTAC 840
 85 CAAACCCATC AGTGAGGTGC TGTCAGAAAT GACAGGCAAC AAGTGGGAT ACACCTTTGA 900
 AGTTATTGGC CATCTTGAAA CCATGATTGA TGCCCTGGCA TCCTGCCACA TGAACATATG 960
 GACCAGCGTG GTTGTAGGAG TTCTCTCATC AGCCAAGATG CTCACCTATG ACCCGATGTT 1020

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

GCTCTTCACT GGAGCCACAT GGAAGGGATG TGTCTTTGGA GGTITGAAAA GCAGAGATGA 1080
TGTCCCAAAA CTAGTGACTG AGTTCTGGGC AAAGAAATTT GACCTGGACC AGTTGATAAC 1140
TCATGTTTAA CCAATTTAAA AAATCAGTGA AGGATTGAGC CTGCTCAATT CAGGACAAAAG 1200
CAITCGAACG GTCCTGACGT TTTGAGATCC AAAGTGGCAG GAGGTCTGTG TTGTCACTGT 1260
GAACCTGGAGT TTCTCTTGTG AGAGTTCCCT CATCTGAAAT CATGTATCTG TCTCACAAT 1320
ACAAGCATAA GTAGAAGATT TGTGGAAGAC ATAGAACCCT TATAAGAAT TATTAAACCTT 1380
TATAACATT TAAAGTCTTG TGAGCACCTG GGAATTAGTA TAATAACAAT GTTAATATTT 1440
TTGATTTACA TTTTGTAAAG CTATAATTGT ATCTTTTAAAG AAAACATACA CTTGGATTTC 1500
TATGTTGAAA TGGAGATTTT TAAGAGTTT AACCAGCTGC TGCAGATATA TAACTCAAAA 1560
CAGATATAGC GTATAAAGAT ATAGTAAATG CATCTCCAG AGTAATATTC ACTTAACACA 1620
TTGAACTAT TATTTTGTAG ATTTGAATAT AAATGTATTT TTTAAACACT TGTATGAGT 1680
TAACTGGAT TACATTTTGA AATCAGTTCA TTCCATGATG CATATTACTG GATTAGATTA 1740
AGAAAGACAG AAAAGATTAA GGGACGGGCA CATTTTTCAA CGATTAAAGAA TCATCATTAC 1800
ATAACTGGT GAAAGTGAAG AAGTATATCA TATGGGTACA CAAGGCTATT TGCCAGCATA 1860
TATTAAATAT TTAGAAAATA TTCCTTTTGT AATACTGAAT ATAAACATAG AGCTAGAGTC 1920
ATATTATCAT ACTTATCATA ATGTTCAAAT TGATACAGTA GAATTGCAAG TCCCTAAGTC 1980
CCTATTCACT GTGCTTAGTA GTGACTCCAT TTAATAAAAA GTGTTTGTAG TTTTAAACAA 2040
CTAAACCG

Seq ID NO: 101 Protein sequence:
Protein Accession #: NP_000664

1 11 21 31 41 51
MGTAGKVIKC KAAVLWEQKQ PPSIEEIEVA PPKTKEVRIK ILATGICRTD DHVIRKTMVS 60
KFPVIVGHEA TGVIESIGEG VTTVKPGDKV IPLFLPQCRE CNACRNPNGN LCIRSDITGR 120
GVLDAGTTRF TCKGKPVHVF MNTSTPTEYT VVDESSVAKI DDAAPPEKVC LIGCGFSTGY 180
GAAVKTGKVK PGSTCVFVGL GGVGLSVIMG CKSAGASRII GIDLNKDKFE KAMAVGATEC 240
ISPKDSTKPI SEVLSEMTGN NVGYTFEIVG HLETMIDALA SCHMNYGTSV VVGVPSPAKM 300
LTYDPHLLFT GRTWKGCVFG GLKSRDDVFK LVTEFLAKKF DLDQLITHVL PFKKISEGFE 360
LLNSGQSIRT VLTF

Seq ID NO: 102 DNA sequence
Nucleic Acid Accession #: NM_006783.1
Coding sequence: 1..786

1 11 21 31 41 51
ATGGATTGGG GGACGCTGCA CACTTTCATC GGGGGTGTCA ACAACACTC CACCAGCATC 60
GGGAAGGTGT GGATCACAGT CATCTTTATT TTCGAGTCA TGATCCTAGT GGTGGCTGCC 120
CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTCGTCTGCA ACACACTGCA ACCGGGATGC 180
AAAAATGTGT GCTATGACCA CTTTTTCCCG GTGTCCCA CA TCCGGCTGTG GGCCCTCCAG 240
CTGATCTTCT TCTCCACCCC AGCGCTGCTG GTGGCCATGC ATGTGCGCTA CTACAGGCAC 300
GAAACCACTC GCAAGTTTCAG GCGAGGAGAG AAGAGGAATG ATTTCAAAGA CATAGAGGAC 360
ATTAATAAAGC ACAGAGTTG GATAGAGGGG TCGCTGTGGT GGACGTACAC CAGCAGCATC 420
TTTTTCGAAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTTACTTCCT TTCAATGGG 480
TACCACCTGC CCTGGGTGTT GAAATGTGGG ATTGACCCCT GCCCAACCT TGTGACTGC 540
TTTATTCTTA GGCCCAACAGA GAAGACCGTG TTTACCATTT TTATGATTTT TCGCTCTGTG 600
ATTTCGATGC TGCTTAACGT GGCAGAGTTG TGCTACCTGC TGCTGAAGT GTGTTTTAGG 660
AGATCAAGAG GAGCAGACAG GCAAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG 720
CAGAAAGAAA TGAATGAGCT GATTTCAGAT AGTGGTCAAA ATGCAATCAC AGGTTTCCCA 780
AGCTAA

Seq ID NO: 103 Protein sequence:
Protein Accession #: NP_006774.1

1 11 21 31 41 51
MDWGLHTFI GGVNKHSTSI GKVNITVIFI FRVMIIVVAA QEVNGDEQED FVCNTLQPGC 60
KNVCYDHPFP VSHIRLWALQ LIPVSTPALL VAMHVAYYRH ETTRKFRRGE KRNDFRDIED 120
IKKHVRIEG SLNWTYTSII FFRIIFEAFF MYVPYFLYNG YHLPWVLKCG IDPCPNLVDC 180
PISRPTKTV FTIFMISASV ICMLLNVAEL CYLLLVKCFR RSKRAQTQKN HPNHALKESK 240
QNEMNELISD SQQNATTFPP S

Seq ID NO: 104 DNA sequence
Nucleic Acid Accession #: NM_020411
Coding sequence: 86-526

1 11 21 31 41 51
GGACCTGGGA AGGAGCATAG GACAGGGCAA GGCGGGATAA GGAGGGGCAC CACAGCCCTT 60
AAGGCAAGAG GGAACCTCAC TGCGCATGCT CTTTGGTGC CCACCTCAGT GCGCATGTTT 120
ACTGGGCGTC TTCCCATCGG CCCCTTCGCC AGTGTGGGGA ACCGCGCGGA GCTGTGAGCC 180
GGCGACTCGG GTCCCTGAGG TCTGGATTCT TTCTCGCTA CTGAGACAGC GCGGACACAC 240
ACAAACACAG AACCACACAG CCACTCCAG GAGCCAGTA ATGGAGAGCC CCAAAAAGAA 300
GAACCAAGAG CTGAAAGTGC GGATCTTACA CCTGGGCAGC AGACAGAAGA AGATCAGGAT 360
ACAGCTGAGA TCCAGTGGG CGACATGGAA GGTGATCTGC AAGAGCTGCA TCAGTCAAAC 420
ACCGGGGATA AATCTGGATT TGGGTTCCGG CGTCAAGGTG AAGATAATAC CTAAAGAGGA 480
ACACTGATAA ATGCCAGAG CAGGTGAAGA GCAACACAA GTTTAAATGA AGACAAGCTG 540
AAACAACGCA AGCTGGTTTT ATATTAGATA TTTGACTTAA ACTATCTCAA TAAAGTTTTG 600
CAGCTTTCAC CAAAAA AAAA

Seq ID NO: 105 Protein sequence:
Protein Accession #: NP_065144.1

1 11 21 31 41 51

MLLWCPPOCA CSLGVFPSPAP SPVWGTTRSC EPATRVPEVW ILSPLLRHGG HTQTQNHAS 60
PRSPVMESEPK KKNQQLKVG I LHLGSRQKKI RIQLRSQCAT WKVICKSCIS QTPGINLDLG 120
SGVKVKIIPK EECCKMPEAG EBPQOV

Seq ID NO: 106 DNA sequence
Nucleic Acid Accession #: J04129
Coding sequence: 99-587

1 11 21 31 41 51
CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCGC AGCCATGCTG TGCCTCCTGC 60
TCACCTCTGG CGTGCCCTGT GTCTGTGGTG TCCCGGCCAT GGACATCCCC CAGACCAAGC 120
AGGACCTGGA GCTCCCAAG TTGGCAGGGA CTTGGCAGTC CATGGCCATG GCGACCAACA 180
ACATCTCCCT CATGGCGACA CTGAAGGCC CTCTGAGGT CCACATCACC TCACTGTGTC 240
CCACCCCCGA GGACAACTG GAGATCGTTC TGACAGATG GGAGAACAA AGCTGTGTTG 300
AGAAGAAGGT CTTGGAGAG AAGACTGGGA ATCCAAAGAA GTTCAAGAT AACTATAOGG 360
TGGCAACGA GGCCAGCTG CTGATACTG ACTACGACAA TTCTCTGTT CTCTGCCTAC 420
AGGACACCAC CACCCCATC CAGAGCATGA TGTGCCAGTA CTTGGCCAGA GTCTCTGTTG 480
AGGACGATGA GATCATGAG GGATTCATCA GGGCTTTGAG GCCCTTGCCC AGGCACCTAT 540
GGTACTTGCT GGACTTGAAA CAGATGGAAG AGCGGTGCGG TTCTAGCTC ACCTCGGCT 600
CCAGGAAGAC CAGACTCCCA CCCTTCCACA CTTCCAGAGC AGTGGGACTT CCTCTGCCC 660
TTTCAAGAA TAACCAAGC TCAGAAGAC ATGACGTGGT CATCTGTGTC GCCATCCCT 720
TCCTGCTGCA CACTGCAAC ATTGCCATGG GGAGGCTGCT CCTGGGGGC AGAGTCTCTG 780
GCAGAGTTA TTAATAAAC CTTGGAGCAT G

Seq ID NO: 107 Protein sequence:
Protein Accession #: AAA60147

1 11 21 31 41 51
MDIPQTKQDL ELPKLAGTWH SMAMATWNIS LMATLKAPLR VHITSLPTP EDNLEIVLHR 60
WENNSCVKVK VLGEKTNPK KFKINYTVAN EATLDDTYD NFLFLCLQDT TTPIQSMQCQ 120
YLARVLVEDD EIMQGFIRAF RPLPRHLWYL LDLKQMEPC RF

Seq ID NO: 108 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 48-794

1 11 21 31 41 51
TCCAGGCAG CAGTTAGCCC GCGGCCCGCC TGTGTGTCCC CAGAGCCATG GAGAGAGCCA 60
GTCTGATCCA GAAGGCCAAG CTGGCAGAGC AGGCCGAACG CTATGAGGAC ATGGCAGCCT 120
TCATGAAGAG CGCCGTGGAG AAGGGCGAGG AGCTCTCCTG CGAAGAGCGA AACCTGCTCT 180
CAGTAGCCTA TAAGAACGTG GTGGGCGGCC AGAGGGCTGC CTGGAGGGTG CTGTCCAGTA 240
TTGAGCAGAA AAGCAACGAG GAGGGCTCGG AGGAGAAGGG GCCCGAGGTG CGTGAATACC 300
GGGAGAAGGT GGAGACTGAG CTCCAGGCGG TGTGCGACAC CGTGTGCGG CTGTGGACA 360
GCCACCTCAT CAAGGAGGCC GGGGACGCGG AGAGCCGGT CTCTACCTG AAGATGAAG 420
GTGACTACTA CGCTCACTG GCGAGGTGG CCACGGTGA CGACAAGAAG CGCATCATG 480
ACTCAGCCCG GTACGCTAC CAGGAGGCCA TGGACATCAG CAAGAAGGAG ATGCCGCCCA 540
CCAACCCCAT CGGCTGGGC CTGGCCCTGA ACTTTCCGT CTCCACTAC GAGATCGCCA 600
ACAGCCCCGA GGAGGCCATC TCTCTGCCA AGACCACTTT CGACGAGGCC ATGGCTGATC 660
TGCACACCT CAGCGAGGAC TCCTACAAAG ACAGCACCTT CATCATGCAG CTGCTCGAG 720
ACAACCTGAC ACTGTGAGC GCGACAAAG CCGGGGAAGA GGGGGCGAG GCTCCCCAGG 780
AGCCCCAGAG CTGAGTGTG CCGGCCACCG CCGGCCCTG CCGCTCCAG TCCCCACCC 840
TGCCGAGAG ACTAGTATG GGTGGAGGCC CCCACCTTC TCCCTAGGC GCTGTCTTG 900
CTCCAAAGGG CTCGTGGAG AGGGACTGGC AGAGCTGAG CCACCTGGG CTGGGATCC 960
CACTCTCTT CGAGCTGTG AGCGACCTA ACCACTGGT ATGCCCCAC CCCTGCTCTC 1020
CGACCCGCT TCCTCCGAC CCCAGGACCA GGCTACTTCT CCGCTCTCT TGCCTCCCTC 1080
CTGCCCTGC TGCCTCTGAT CGTAGGAATT GAGGAGTGT CCGCTTGTG GCTGAGAACT 1140
GGACAGTGGC AGGGGCTGAT GATGGGTGT TGTGTGTGT TGTGTGTGT TGTGTGTGT 1200
CGCGCGGCC AGTGCAAGAC CGAGATTGAG GAAAGCATG TCTGTGGT GTGACCATGT 1260
TTCCTCTCAA TAAAGTTCCC CTGTGACACT C

Seq ID NO: 109 Protein sequence:
Protein Accession #: NP_006133.1

1 11 21 31 41 51
MERASLIQKA KLAQAERYE DMAAFMKGAV EKGEELSCEE RNLLSVAYKN VVGQRAAWR 60
VLSSIEQKSN EEGSEKGP VREYREKVT ELQGVCDTVL GLLDHSLIKE AGDAESRVFY 120
LKMKGDIYRY LAEVAETDDK KRIIDSARSA YQEAIDISK EMPPTNPIRL GLALNFSVFH 180
YEIANSPEEA ISLAKTTTDE AMADLHTLSE DSYKDSTLIM QLLRDLNLTW TADNAGEEGG 240
EAPQEPQS

Seq ID NO: 110 DNA sequence
Nucleic Acid Accession #: NM_000695
Coding sequence: 407-1564

1 11 21 31 41 51
CACGAGTTGG TTTGGGAGCT GCCAGTCTCC TGGGAGGATC GCAGTCAGCA GAGCAGGGCT 60
GAGGCCTGGG GGTAGGAGCA GAGCCTGCGC ATCTGGAGGC AGCATGTCCA AGAAAGGGAG 120
TGGAGGTGCA GCGAAGGACC CAGGGGCGA GCCACGCTG GGGATGAGC CCTTCGAGGA 180
CACACTGCGG CGGCTGCGT AGGCCTTCAA CTGAGGGCGC ACGCGGCGG CCGAGTCCG 240
GGCTGCGCAG CTCACGGGC TGGGCCACTT CCTCAAGAA AACAGCAGC TTCTGCGCGA 300

5	CGTGCTGGCC	CAGGACCTGC	ATAAGCCAGC	TTTCGAGSCA	GACATATCTG	AGCTCATCCT	360
	TTGCCAGAAC	GAGGTGTACT	AGCTCTCTCA	GAACCTTCAG	GCCTGGATGA	AGGATGAACC	420
	ACGGTCCACG	AACCTGTTC	TGAAGCTGGA	CTCGGTCTTC	ATCTGGAAGG	AACCTTTGG	480
	CTCGGTCTTC	ATCATCGCAC	CCTGGAACTA	CCCATTTGAC	CTGACCCCTG	TGCTCCTGGT	540
	GGGCACCTCC	CCCGCAGGGA	ATTGGGTGGT	GCTGAAGCCG	TCAGAAATCA	GCCAGGGCAC	600
	AGAGAAGGTC	CTGGCTGAGG	TGCTGCCCCA	GTAACCTGAC	CAGAGCTGCT	TTGCCGTGGT	660
	GCTGGGCGGA	CCCCAGGAGA	CAGGGCAGCT	GCTAGAGCAC	AAGTTGGACT	ACATCTTCTT	720
	CACAGGGAGC	CCTCGTGTGG	GCAAGATTGT	CATGACTGCT	GCCACCAAGC	ACCTGAGCCC	780
	TGTCACCTCG	GAGCTGGGGG	GCAAGAACCC	CTGCTACGTG	GACGACAACT	GCGACCCCCA	840
10	GACCGTGGCC	AACCGCTGGG	CCTGGTTCTG	CTACTTCAAT	GCCGGCCAGA	CCTGCGTGGC	900
	CCCTGACTAC	GTCTGTGTCA	GCCCCGAGAT	GCAGGAGAGG	CTGCTGCCCG	CCCTGCAGAG	960
	CACCATCACC	CGTTTCTATG	GCGAGACCCC	CCAGAGCTCC	CCAAACCTGG	GCCGCATCAT	1020
	CAACCCAGAAA	CAGTTCACAG	GGCTGCGGGC	ATTGCTGGGC	TGCGGCCGCG	TGGCCATTGG	1080
	GGGCCAGAGC	AACGAGAGCG	ATCGCTACAT	CGCCCCCAGC	GTGCTGGTGG	ACGTGCAGGA	1140
15	GACCGAGCCT	GTGATGTCAG	AGGAGATCTT	CGGGCCCATC	CTGCCCATCG	TGAACGTGCA	1200
	GAGCGTGGAC	GAGGCCATCA	AGTTCATCAA	COGGCAGGAG	AAGCCCCCTG	CCCTGTACGC	1260
	CTTCTCCAAC	AGCAGACAGG	TTGTGAACCA	GATGCTGGAG	CGGACCAAGC	GCCGAGCTT	1320
	TGGAGGCAAT	GAGGGCTTCA	CCTACATATC	TCTGCTGTCC	GTGCCATTGG	GGGGAGTCCG	1380
20	CCACAGTGGG	ATGGGCGGCT	ACCACGGCAA	GTTCACTTTC	GACACCTTCT	CCCAACCCCG	1440
	CACCTGCTCT	CTCGCCCCCT	COGGCTTGGG	GAAATTAAGG	GAGATCCGCT	ACCCACCCCTA	1500
	TACCGACTGG	AACCAGCAGC	TGTTACGCTG	GGGCATGGGC	TCCAGAGCT	GCACCCCTCT	1560
	GTGAGCGTCC	CACCGCGCTC	CAACGGGTCA	CACAGAGAAA	CCTGAGTCTA	GCCATGAGGG	1620
	GCTTATGCTC	CCAACTCACA	TTGTTCTCTC	AGACCGCAGG	CTCCCCCAGC	CTCAGGTTCG	1680
25	TGGAGCTGTC	ACATGACTGC	ATCCTGCTCG	CCAGGGCTGC	AAAGCAAGGT	CTTGCTTCTA	1740
	TCTGGGGGAC	GCTGCTCGAG	AGAGGCGGAG	AGGCGGCAGA	ACATGCCAGG	TGTCCTCACT	1800
	CACCCCAACC	TCCCAATTC	CAGCCCTTTC	CCCTCTCGST	CAGGGTTGGC	CAGGCCACGT	1860
	CACAGGGGCA	GTGTACCCCT	GGAAAATACA	GTGCCCTGCC	TTCTTAGGGG	CATCAGCCCT	1920
	GAAAGGTGTA	GAGCGTGGAG	CCCTCCAGGC	CTTTGCTCTC	CCCTCTAGGC	ACACGCGCAC	1980
30	TTCCACCTCT	GCCCATCCCC	AACGACCA	GCATGCTCTC	CCCAAGGGAT	CCTCTCATAT	2040
	CCCACACTGG	TCTCTGCACC	ACCCCTCTGG	TTACACCCGC	ACCCTGCACT	CACCCACAGC	2100
	AGCTCCATCC	GTGGGAAAA	CTGGGGTTTG	CATCACTCCA	CTGCACAGTG	TTAGTGGGAC	2160
	CTGGGGGCAA	GTCCCTTGAC	TTCTCTGAGC	CTCAGTTTCC	TTATGTGAAA	GTTGCTGGAA	2220
	CCAAAATGGA	GTCACTTATG	CCAAACTCTA	ATAAAATGGA	CTCGGGGGGG	CACATAGAAG	2280
35	CCCTCACACA	CACATGCCCG	TAAACAGATT	TATCACCAAG	ACACGCTGCG	ATGTAAGACC	2340
	AGACACAGGG	CGTATGGA	AGCACGTCTC	CAAGAGCTGT	AGTATTCCAG	ATGAGCTGCA	2400
	GATGCTTACC	TACCACGGCC	GTCTCCACCA	GAAGAACATC	GCCAACTCCT	GCGATCAGCT	2460
	TGTGACTTAC	AAACCTTGTT	TAAAGCTGTC	TTACATGGAC	TTCTGTCTCT	TAAACGTTTC	2520
40	CCCTTGCTGT	TGGCCCTCTG	TGTATGCTGT	GGATCCTTCC	AAGCACTCAT	AGCCACAGATA	2580
	GGAATCCTCT	GCTCTCTCCA	AATAAATTC	TCTGTTC			

Seq ID NO: 111 Protein sequence:
Protein Accession #: NP_000686

45	1	11	21	31	41	51	
	1	11	21	31	41	51	
	MDPEPRSTNL	FMKLDSPFIW	KEPFGVLVII	APWNYPLNLT	LVLLVGLTLP	GNCVVLKPS	60
	ISQTEKVL	EVLPQYLDQS	CFAVVLGGPQ	ETGQLLEHL	DYIFFTGSPR	VGRIVMTAAT	120
	KHLTPVTL	GGKAPFCVDD	NCDPQTVANR	VAWFCYFNAG	QTCVAPDYVL	CSPMEQERLL	180
50	PALQSTITR	YGDPDFQSPN	LGRIINQKQF	QRLRALLGCG	RVAIGGGSNE	SDRYIAPTUL	240
	VDVQETEPVM	QEEIFGPILP	IVNVQSVDEA	IKPINRQEPK	LALYAFNSNR	QVVNQMLERT	300
	SSGSFGNBE	PTYISLISVF	FGGVGSHGMS	RYHGKPTFTD	FSHRTCLLA	PSGLEKLKEI	360
	RYPPYTDWNG	QLLRWNGMSQ	SCTLL				

Seq ID NO: 112 DNA sequence
Nucleic Acid Accession #: NM_004456
Coding sequence: 58-2298

60	1	11	21	31	41	51	
	1	11	21	31	41	51	
	GAATTCGGGG	CGACGCGCGG	GAACAACCGC	AGTCGGGCGG	CGGACGAAG	AATAATCATG	60
	GGCCAGACTG	GGAGAAATC	TGAGAAGGGA	CCAGTTTGT	GGCGGAAGCG	TGTAAATCA	120
	GAGTACATGC	GACTGAGACA	GCTCAAGAGG	TTCAGACGAG	CTGATGAAGT	AAAGAGTATG	180
	TTAGTTCCA	ATCGTCAGAA	AATTTTGAA	AGAACGGAAA	TCTTAAACCA	AGAAATGAAA	240
65	CAGCGAAGGA	TACAGCCTGT	GCACATCTCG	ACTTCTGTGA	GCTCATTGCG	CGGGACTAGG	300
	GAGTGTTCGG	TGACCACTGA	CTTGATTTT	CCAACACAAG	TCATCCCAT	AAAGACTCTG	360
	AATGCAGTTG	CTTCAGTACC	CATAATGTAT	TCTTGGTCTC	CCCTACAGCA	GAATTTTATG	420
	GTGGAAGATG	AACTGTGTTT	ACATAACATT	CCTTATATGG	GAGATGAAGT	TTTAGATCAG	480
	GATGGTACTT	TCAATTGAGA	ACTAATAAAA	AATTATGATG	GGAAAGTACA	CGGGATAGA	540
70	GAATGTGGGT	TTATAAATGA	TGAAATTTT	GTGGAGTTGG	TGAATGCCCT	TGGTCAATAT	600
	AATGATGATG	ACGATGATGA	TGATGGAGAC	GATCCTGAAG	AAAGAGAAGA	AAAGCAGAAA	660
	GATCTGGAGG	ATCACCAGAG	GATATAAGAA	AGCCGCCAC	CTCGGAAAT	TCCTTCTGAT	720
	AAAAATTTGG	AGGCCATTTC	CTCAATGTTT	CCAGATAAGG	GCACAGCAGA	AGAACTAAAG	780
	GAAAAATATA	AGAACTCAC	CGAACAGCAG	CTCCAGGCG	CACCTTCTCC	TGAATGTACC	840
75	CCCAACATAG	ATCGACCAAA	TGCTAAATCT	GTTCAAGAG	AGCAAAAGCTT	ACACTCCTTT	900
	CATACGCTTT	TGTAGGCGG	ATGTTTTAAA	TATGACTGCT	TCCTACATCC	TTTTCATGCA	960
	ACACCCACCA	CTTATAAGCG	GAAGAACACA	GAACAGCTC	TAGACAACAA	ACCTTGTGGA	1020
	CCACAGTGT	AACTCTCTGA	GGAGGGAGCA	AAGGAGTTTG	CTGCTGCTCT	CACCGCTGAG	1080
	CGGATAAGAA	CCCCACCAAA	ACGTCCAGGA	GGCCGCGAAA	GAGGACGGCT	TCCCAATAAC	1140
80	AGTAGCAGGC	CCAGCACCCC	CACCATTAAT	GTGCTGGAAT	CAAAGGATAC	AGACAGTGAT	1200
	AGGGAAGCAG	GGACTGAAAC	GGGGGGAGAG	AACAATGATA	AAGAAGAAGA	AGAGAAGAAA	1260
	GATGAAACTT	CGAGCTCCTC	TGAAGCAAA	TCTCGGTGTC	AAACACCAAT	AAAGATGAAG	1320
	CCAAATATTG	AACTCTCTGA	GAATGTGGAG	TGGAGTGGTG	CTGAAGCTCT	AATGTTTAGA	1380
	GTCTCTATTG	GCATCTACTA	TGCAATTTT	TGTGCCATTG	CTAGGTTAAT	TGGGACCAAA	1440
85	ACATGTAGAC	AGGTGTATGA	GTTTAGAGTC	AAAGAATCTA	GCATCATAGC	TCCAGCTCCC	1500
	CGTAGAGATG	TGGATATCTC	TCCAAGGAAA	AAGAAGAGGA	AACACCGGTT	GTGGGCTGCA	1560
	CATGCGAGAA	AGATACAGCT	GAAAAAGGAC	GGCTCCTCTA	ACCATGTTTA	CAACTATCAA	1620

5
 10
 15

```

CCCTGTGATC ATCCACGGCA GCCITGTGAC AGTTCGTGCC CTGTGTGTGAT AGCACAATAAT 1680
TTTTGTGAAA AGTTTTGTGCA ATGTAGTICA GAGTGTCAAA ACCGCTTTCC GGGATGCCGC 1740
TGCAAGACAC AGTGCACAC CAAGCAGTGC CGTGCTACC TGGCTGTCCG AGAGTGTGAC 1800
CCTGACCTCT GTCTTACTTG TGGAGCCGCT GACCATGGG ACAGTAAAAA TGTGTCTCTG 1860
AAGAACTGCA GTATTACGG GGGCTCCAAA AAGCATCTAT TGTGGCACC ATCTGACGTG 1920
GCAGGCTGGG GGAATTTTAT CAAAGATCCT GTGCAGAAAA ATGAATTCAT CTCAGAATAC 1980
TGTGGAGAGA TTATTTCTCA AGATGAAGCT GACAGAAGAG GGAAGTGTG TGATAAATAC 2040
ATGTGCAGCT TTCTGTTCAA CTGAAACAAT GATTTGTGG TGGATGCAC CCGCAAGGGT 2100
AACAAATTC GTTTTGCAAA TCATTGCGTA AATCCAACT GCTATGCAAA AGTTATGATG 2160
GTTAACGGT ATCAGAGAT AGGTATTTT GCCAAGAGAG CCATCCAGAC TGGCGAAGAG 2220
CTGTTTGTG ATTACAGATA CAGCCAGGCT GATGCCCTGA AGTATGTCG CATCGAAAGA 2280
GAAATGGAAA TCCCTTGACA TCTGCTACCT CCTCCCCCTC CTCTGAAACA GCTGCCCTAG 2340
CTTCAGGAAC CTCGAGTACT GTGGGCAATT TAGAAAAAGA ACATGCAGTT TGAAATTTCTG 2400
AATTGCAAA GTACTGTAAG AATAATTTAT AGTAATGAGT TTAATAATCA ACTTTTATT 2460
GCCTTCTCAC CAGCTGCAAA GTGTTTGTG CAGTGAATT TTTGCAATAA TGCAATATGG 2520
TACATTTTTC AACTTTGAAT AAAGAATACT TGAACCTGAA AAAAAAATA AAAAAA
  
```

20
 Seq ID NO: 113 Protein sequence:
 Protein Accession #: NP_004447

25
 30
 35

```

1 11 21 31 41 51
MGQTGKKSEK GPVCRWRKVK SEVMRLRQLK RFRRADDEVKS MFSSNRQKIL ERTEILNQEW 60
KQRRIQPVHI LTSVSSLRGT RECVTSDDL PFTQVILPKT LNAVASVPIM YSWSPLOQNF 120
MVDETVLHN IPYMGDEVLQ ODGTFIEELI KNYDGKVHGD RECGFINDEI FVELVNALGQ 180
YNDDDDDDDG DDPEEREERQ KDLEDHRDDK ESRPPRKFPK DKILEAISSM PPDKGTAEL 240
REKYKELTQ QLPALPPEC TPNIDGPNK SVQREQSLHS FHTLPCRRCF KYDCFLHPFH 300
ATPNTYKRN TETALDNKPC GPQCYHLEB AKEFAAALTA ERIKTPPKRP GRRRRGRLPN 360
NSSRPSTPTI NVLESKDTDS DREAGTETGG ENNDKEEEK KDSTSSSEA NSRCQPIKM 420
KPNIEPPENV EWSGAESAMP RVLIGTYND FCAIARLIGT KTCRQVYEFK VKESSIIAPA 480
PAEDVDTPPR KKRKRHLWA AHCRIQLKK DGSSNHVYNY QPCDHPQPC DSSCPVIAQ 540
NFCEKFCQS SEQNRFPGC RCKAQCNKQ CPCYLAVREC DPDLCLTCGA ADHWDKSNVS 600
CKNCSIQRGS KKHLLAPSD VAGWGFIDK PVQKNEFISE YCGHISQDE ADRRGKVYDK 660
YMCSLFLNL NDFVVDATRK GNKIRFANHS VNPNCYAKVM MVNGDHRIGI FAKRAQTGE 720
ELFVDYRSQ ADALKVGLB REMEIP
  
```

40
 Seq ID NO: 114 DNA sequence
 Nucleic Acid Accession #: NM_001827
 Coding sequence: 96-335

45
 50
 55

```

1 11 21 31 41 51
AGTCTCCGCG GAGTGTGTC CTGGGCTGGA CGTGGTTTG TCTGCTGCGC CGGCTCTCG 60
CGCTCTCGTT TCATTTCTG CAGCGGCCA CGAGGATGGC CCACAGCAG ATCTACTACT 120
CGGACAGTA CTTCGACGAA CACTACGAGT ACCGGCATGT TATGTTACCC AGAGAACTTT 180
CCAAACAAGT ACCTAATACT CATCTGATGT CTGAAGAGGA GTGGAGGAGA CTGGGTGTC 240
AACAGAGTCT AGGCTGGGT CATTACATGA TTCATGAGCC AGAACACAT ATTCTTCTCT 300
TTAGACGACC TCTTCAAAA GATCAACAAA AATGAAGTTT ATCTGGGGAT CGTCAAATCT 360
TTTTCAAAAT TAATGTATAT GTGTATATAA GGTAGTATTC AGTGAATACT TGAGAAATGT 420
ACAAATCTTT CATCCATACC TGTGCATGAG CTGTATTCTT CACAGCAACA GAGCTCAGTT 480
AAATGCAACT GCAAGTAGGT TACTGTAAAG TGTTTAAGAT AAAAGTTCTT CAGTCAGTT 540
TTTCTCTTAA GTGCTGTTT GAGTTTACTG AAACAGTTTA CTTTGTCTCA ATAAAGTTTG 600
TATGTTGCAT TAAAAAATA AAAAAA
  
```

Seq ID NO: 115 Protein sequence:
 Protein Accession #: NP_001818

60

```

1 11 21 31 41 51
MAHKQIYYSD KYFDEHYEYR HVMLPRELSK QVPKTHLMSE EEWRLIGVQQ SLGWVHYMIH 60
EPEPHLLFR RLPKDKQK
  
```

65
 Seq ID NO: 116 DNA sequence
 Nucleic Acid Accession #: CAT cluster

70
 75
 80
 85

```

1 11 21 31 41 51
TCAGACCTCA TGAGTCACTT GGACTCTTGA GCCACCTCTG GGGGTGGAGT CTCTCTCTG 60
GCATCTGGAC CCTTGGTGCT ATCGACGAAG CTTGGGTGGG GCTCTAGCT GCTATGTGCA 120
AGAGGTGTGT TCCAGGGAAA GCCCTATCT CTCTGCAGAG GTCAAGTGAA AGCGACGGCC 180
GCAGCCAACA GAGTTCAAAA TGCAGGCTTG GAAAGTACAG GGGGCTCTGT GGAGGATGGG 240
AAGACTGAT CCACATCCC ACCAGGAAGT TTAGCAGAAC CCCCGCTGTC CAACTGGACC 300
CCTTGAAGG ACCTGGCTCA GGCTGGACCA CCTCTTGAGA GGGAGGAGCT CTGGATTGTA 360
TCAAGAATTC TTTGCTGAGC ATGGTGCCCT ATGCCTATAA TACCAACACT TTGGAGGACC 420
AGTGTGGGAG GATCTCTTGA GCCCAGGAGT TCAAGACTAG CCTGGGCAAC ACAGAGAGAA 480
CCCATCTCTA AATAATAAAT AATAATAAAA TAAAAAATTA GCAGGCGATG GTGGCATGTG 540
CCTGTAGTTC CAGCTACCCA GGAGGCTGAG GCAAGAGGAT GGCCTGGAGCC TGGGATGTTG 600
AGGCTGCAAT GAACCTGTAT TACCCCACTG CACTCCAGCC TGGGCAAAAG AGCGAGAGAA 660
CCTGTCTCAA AATAATAAAT TAATAATAAT CTTATTTTGG AGAATAAAGA GACCTCTGGA 720
TTTGGGTGC CATTTGGGTA GAAAGAAAAA ACGTTTACAC CGAGAAATAG TCTGTGTTGC 780
CCTGAAGGAG CAGAGGGATG CATCGCTGGA GGTGACCTAC AGTTGAAGAA GACTCATTAT 840
GACAGACCTT GTCTTCTCT CTGTGTGAAA GTGTTCCTCT TGCTGCTACT GCTCATGAGA 900
CTCTCCGCC TCCCTGTCCC AGGGAACCAA AGGCTTTTCT ACCACACCTT TTCTTGCCCC 960
CGGCTCCCA TGTCTGCTGT GCCTTTGTAC TCAGCAATTC TTGTTGCTC CATTATCTTC 1020
CAGCGGATA CAGAGTGAAT AGTTAACCACT ACTTAGGTCA AATAGGATCT AAATTTTGT 1080
TCTGCTCCG TGTAAGAGG CCAGTGTGTT TGTGTTGCAA GCAGCCTTG AATAGTAAT 1140
  
```


WO 02/086443

PCT/US02/12476

5
10
15
20
25
30
35
40
45
50
55
60

```
CTTCTCATTT GTTGGGATC TGGCCACCAA GTTCCAGAAT GATACACGGA TCAGTGACGA 1200
AGTTTCATCAG GCTCTCGGAC CTTAGGGCTG TGGGAGAAGG CTTACAGCAGC AGAAGCTGATG 1260
GTGAAGGCTC GGTGTTCTCA TCCTCAACTT TCTTTGCTTC GATCATACAC AAGAATACAT 1320
TTGGAAGGGC AAAAATGAA CACTGTGCTT CATTGCAGCC GTGTTTGTGT ACACAGATGC 1380
ACAGTCTGCT GTGAAGACCT TCTCTCAAGT GGCATTGGGG AGTCCATGCC AGATCATGGT 1440
GCTTTCATGAG AGACTGACAG CTATCAGGGG TTGTGGCACT TAGTGAGGAC TCTCTCCTCC 1500
CAGTGTGTGC TGATGACACA TACACACCTG ACAATAGCTT GAGTCTCTCT TGTTCCTTTT 1560
ACTCTGTAGC CAACATACAC ATGATTAAAA ACCCTTTCTA AATATCTATC ATGTTTCATC 1620
CTGTGTCCAA TGCAGAGTCA GAGCTATTGT TACTTCATTA TTATTTCCAA GCGGAATAGT 1680
TGGCTTTCTT TTGCAAAAAA TAAATTAAGT TTTTGTATGT TGCAAAAAAA AAAAAAAAAT 1740
AAAAA
```

Seq ID NO: 117 DNA sequence

Nucleic Acid Accession #: BC012178.1

Coding sequence: 204-2285

1 11 21 31 41 51

```
CTTCTCTCCC GCGCGCTGG GCGCGGCTC CGCTGCTGT TGCTCCATTG GCGGCTTTTC 60
TGGCGGCTGG CTCCTCTCCG CTGCGGCTG CTCCTCGACC AGGCCTCCTT CTCACCTCA 120
GCGCGGCTGG CCGACCTTTC GCGCACCTC CCGCCCGTCT TGTACTGTCT GCGGTCACCG 180
CAGCGGCTCC GCGCCTGGCC CCGATGGCTC TGTGCAACGG AGACTCCAAG CTGGAGAATG 240
CTGGAGGAGA CCTTAAGGAT GGCCACCACC ACTATGAAGG AGCTGTGTCT ATTCTGGATG 300
CTGTGTCTCA GTACGGGAAA GTCATAGACC GAAGAGTGAG GGAAGTGTTC GTGCAGTCTG 360
AAATTTTCCC CTGGAAGAAA CCAGCATTGT CTATAAAGGA ACAAGGATTC CGTGCTATTA 420
TCATCTCTGG AGGACCTAAT TCTGTGTATG CTGAAGATGC TCCTGTGTTT GATCCAGCAA 480
TATTCATCTAT TGGCAAGCCT GTTCTTGAAA TTTGCTATGG TATGCAGATG ATGAATAAGG 540
TATTTGGAGG TACTGTGCAC AAAAAAAGTG TCAGAGAAGA TGGAGTTTTC AACATTAGTG 600
TGAATAATAC ATGTGTCATT TTCAGGGGCC TTCAGAAAGGA AGAAGTGTGT TTGCTTACAC 660
ATGGAGATAG TGTAGACAAA GTAGCTGATG GATTCAAGGT TGTGGCACGT TCTGGAACAA 720
TAGTAGCAGG CATAGCAAAAT GAATCTAAAA AGTTATATGG AGCAGAGTTC CACCTGAAG 780
TTGGCCTTAC AGAAAAATGA AAAGTAATAC TGAAGAATTT CCTTTATGAT ATAGCTGGAT 840
GCAGTGGAAC CTTACCGCTG CAGAACAGAG AACTTGAGTG TATTCGAGAG ATCAAAGAGA 900
GAGTAGGCAC GTCAAAAGTT TTGGTTTAC TCAGTGGTGG AGTAGACTCA ACAGTTTGTGTA 960
CAGCTTTGCT AAATCTGTCT TTGAACCAAG AACAAGTCAT TGCTGTGCAC ATTGATAATG 1020
GCTTTATGAG AAAACGAGAA AGCCAGTCTG TTGAAGAGGC CCTCAAAAG CTGGGAATTC 1080
AGGTCAAAGT GATAAATGCT GCTCATTCTT TCTACAATGG AACAACAACC CTACCAATAT 1140
CAGATGAAGA TAGAACCCCA CGGAAAAGAA TTAGCAAAAC GTTAATATAT ACCACAAGTC 1200
CTGAAGAGAA AAGAAAAATC ATTGGGGATA CTTTGTGTTA GATTGCAAT GAAGTAATGT 1260
GAGAAATGAA CTGGAACCAA GAGGAGGTTT TCCTTGCCCA AGGTACTTTA CGGCTGTATC 1320
TAATTGAAGT TGATCTCCTT GTTGCAAGTG GCAAAAGCTG ACTCATCAAA ACCCATCAAA 1380
ATGACACAGA GCTCATCAGA AAGTTGAGAG AGGAGGGAAA AGTAATAGAA CCTCTGAAAG 1440
ATTTTCATAA AGATGAAGTG AGAATTTGG CGAGAGAATC TGGACTTCCA GAAGAGTTAG 1500
TTTCAGGCAC TCATTCTTCA GGTCTGGGCC TGGCAATCAG AGTAATATGT GCTGAAGAAC 1560
CTTATATTGT TAAGGACTTT CCTGAAACCA ACAATATTTT GAAATAGTA GCTGATTTT 1620
CTGCAAGTGT TAAAGAGCAA CATACCTAT TACAGAGATG CAAAGCTGCG ACAACAGAA 1680
AGGATCAGGA GAAGCTGATG CAAATTACCA GTCTGCATTC ACTGAATGCC TTCTGTCTGC 1740
CAATTAACAA TGTAGGTGTG CAGGCTGACT GTCTGTTCTA CAGTTACGTG TGTGGAATCT 1800
CCAGTAAGAA TGAACCTGAC TGGGAATCAC TTATTTTCTT GGCTAGGCTT ATACCTCGCA 1860
TGTGTACAAA CGTTAACAGA GTTGTATTAT TATTTGGCCC ACCAGTTAAA GAACCTCCTA 1920
CAGATGTTAC TCCCATCTTC TTGACACAG GGGTGCTCAG TACTTTACGC CAAGCTGATT 1980
TTGAGGCCCA TAACATTCTC AGGAGTCTG GGTATGCTGG GAAATCAGC CAGATGCCGG 2040
TGATTTTGAC ACCATTACAT TTTGATGGG ACCCACTTCA AAAGCAGCCT TCATGCCAGA 2100
GATCTGTGTT TEEDQEKLMQ ITSLSHINAF LLPRTVGVQ GDCRSYSYVC GISSKDEPDW 540
ESLIFLARLI PRMCHNVNRV VYIFGPPVKE PPTDVTPTFL TTGVLSTLRQ ADFAHNILR 600
CTGGCAATGA GATCCCTGTA GAGGTGGTAT TAAAGATGGT CACTGAGATT AAGAAGATTC 2220
CTGGTATTTT TCGAATTAGT TATGACTTAA CATCAAGGCC CCCAGGAAGT ACTGAGTGGG 2280
AGTAATAAAC TTCTGTCTCT ATTA
```

Seq ID NO: 118 Protein sequence:

Protein Accession #: AAH12178.1

1 11 21 31 41 51

```
MALCNGDKSL ENAGGDLKDG HHHYEGAVVI LDAGAQQYGVK IDRRVRELTV QSEIFFLETP 60
APAIKEQGER AIIISGGPNS VVAEDAPWFD PAIFTIGKPV LGICYGMQMM NKVFGGTVHK 120
KSVREDGVFN ISVDNTCSLF RGLQKEEVVL LTHGDSVDKV ADGFKVVARV GNIVAGIANE 180
SKKLYGAQFH PEVGITENGK VILKNFLYDI AGCSGFTTVQ NRELECIRES KERVGTSKVL 240
VLLSGGVDSV VCTALLNRL NQEQVIAVHI DNGFMRKRES QSVSEALKKL GIQVKVINAA 300
HSFYNGTTLT PISDEDRTPR KRISKILNMT TSPEEKRII GDTFVKIANE VIGEMNLKPE 360
EVFLAQGLLR PDLIESASLV ASGKAELIKT HNDTELIRK LREBKVIEP LKDPHKDEVR 420
ILGRELGLPE ELVSRHPPFG PGLAIRVICA EEPYICKDFP ETRNLIKIVA DPSASVKKPH 480
TLLQRVKACT TEEDQEKLMQ ITSLSHINAF LLPRTVGVQ GDCRSYSYVC GISSKDEPDW 540
ESLIFLARLI PRMCHNVNRV VYIFGPPVKE PPTDVTPTFL TTGVLSTLRQ ADFAHNILR 600
ESGYAGKISQ MPVILTPLEH DRDPLQKQPS QRSVVIRTG ITSDFMTGIP ATPGNEIPVE 660
VVLKMTVEIK KIPGISRIMY DLTSKPPGTT EWE
```

Seq ID NO: 119 DNA sequence

Nucleic Acid Accession #: NM_006500.1

Coding sequence: 27..1967

1 11 21 31 41 51

```
ACTTGCGTCT CCGCTCTCGG CCAAGCATGG GGCTTCCAG GCTGGTCTGC GCCTTCTTGC 60
TGCGCGCTCT CTGCTGCTGT CCTCGCTGCG CGGTGTGGCC CGGAGAGGCT GAGCAGCTCT 120
CGCTGAGCTT GGTGGAGGTG GAAGTGGGCA GCACAGCCTT TCTGAAGTGC GGCTCTCTCC 180
AGTCCCAAGG CAACCTCAGC CATGTCGACT GGTTTTCTGT CCACAAGGAG AAGCGGAGC 240
```


5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

TCACTCTCCG TGTGGCGCAG GCCCAGGGCC AGAGCGAACC TGGGGAGTAC GAGCAGGGCC 300
 TCAGCCCTCA GGAAGAGAGG GCTACTCTGG COCTGACTCA AGTCACCCCC CAAGACGAGC 360
 GCACTCTCTT GTGCCAGGGC AAGCGCCCTC GGTCCCAGGA GTACCGCATC CAGCTCCGGC 420
 TCTACAAAGC TCCCGAGGAG CCAACATCC AGGTCAACCC CCTGGGCATC CCTGTGAACA 480
 GTAAGGAGCC TGAGGAGGTC GCTACCTGTG TAGGGAGGAA CGGTACCCCC ATTCTCAAG 540
 TCATCTGGTA CAAGAATGCG CGGCCCTCTG AGGAGGAGAA GAACCGGGTC CACATTCAAGT 600
 CGTCCAGAC TGTGGAGTGG AGTGGTTTGT ACACCTTGCA GAGTATTCTG AAGGCACAGC 660
 TGGTTAAAGA AGACAAAGAT GCCCAGTTTT ACTGTGAGCT CAACTACCGG CTGCCCATGT 720
 GGAACCATCT GAAGGAGTCC AGGGAAGTCA CGTCCCTGT TTTCTACCGG ACAGAAAAG 780
 TGTGGCTGGA AGTGGAGCCC GTGGGAATGC TGAAGGAAGG GGACCGCGTG GAAATCAGGT 840
 GTTTGGCTGA TGGCAACCTC CCACCACTAC TCAGCATCAG CAAGCAGAAC CCCAGCACCA 900
 GGGAGGCAGA GGAAGAGACA ACCAACGACA ACAGGGGTCT GGTCTGGAG CCTGCCCGA 960
 AGGAACACAG TGGCGCTAT GAATGTGAGG CCTGGAACCT GGACACCATG ATATCGCTGC 1020
 TGAGTGAACC ACAGGAACCT CTGGTGAACCT ATGTGTCTGA CGTCCGAGTG AGTCCCGCAG 1080
 CCCCTGAGAG ACAGGAAGGC AGCAGCCTCA COCTGACCTG TGAGGCAGAG AGTAGCCAGG 1140
 ACCTCGAGTT CCAGTGGCTG AGAGAAGAGA CAGACCAAGT GCTGGAAAGG GGGCCTGTGC 1200
 TTAGTTGCA TGACCTGAAA CGGGAGGCAG GAGGCGGCTA TCGCTGCGTG GCGTCTGTGC 1260
 CCAGCATACC CGGCCTGAAC CGCACACAGC TGTCAAGCT GGCCATTTTT GGCCCCCTT 1320
 GGTATGGCATT CAAGGAGAGG AAGGTGTGGG TGAAGAGAA TATGGTGTG AATCTGTCTT 1380
 GTGAAGCGTC AGGGCACCCC CGGCCACCA TCTCTGGAA CGTCAACGGC ACGCCAAGTG 1440
 AACAAAGACA AGATCCACAG CGAGTCTCTG GCAACCTGAA TGTCTCTGTG ACCCCGAGGC 1500
 TGTGTGAGAC AGGTGTGAAA TGCAACGCTT CCAACGACCT GGGCAAAAC ACCAGCATCC 1560
 TCTTCTGGA GCTGGTCAAT TTAACCAACC TCACACAGA CTCACACA ACCACTGGCC 1620
 TCAGCACTTC CACTGCCAGT CCTCATACCA GAGCAACAG CACCTCCACA GAGAGAAAGC 1680
 TGCCGGAGCC GGAGAGCCGG GCGGTGGTCA TCGTGGCTGT GATTGTGTGC ATCTCTGGTCC 1740
 TGGCGGTGCT GGGCGCTGTC CTCTATTTC TCTATAAGAA GGGCAAGCTG CCGTGCAGGC 1800
 GCTCAGGGA GCAAGAGATG ACCTGCCCCG CGTCTGTA GACCGAATT GTAGTTGAAG 1860
 TTAAGTCAGA TAAGCTCCCA GAAGAGATGG GCCTCTGCA GGGCAGCAGC GGTGACAAGA 1920
 GGGCTCCGG AGACAGGGA GAGAAATACA TCGATCTGAG GCATTAGCCC GGAATCACTT 1980
 CAGCTCCCTT CCTGCTGAG ACCATTCCCA GCTCCCTGCT CACTCTTCTC TCAGCCAAAG 2040
 CCTCCAAAG GACTAGAGAG AAGCCTCTCT CTCCTCTCAC CTGCACACC CCTTTAGAG 2100
 GGCCACTGGG TTAGGACCTG AGGACCTCAC TTGGCCCTGC AAGCCGCTTT TCAGGGACCA 2160
 GTCCACCACC ATCTCTCCA CGTTGAGTGA AGCTCATCCC AAGCAAGGAG CCCAGTCTC 2220
 CCGAGCGGGT AGGAGAGTTT CTTCGAGAAC GTGTTTTTC TTTACACACA TTATGGCTGT 2280
 AAATACCTGG CTCTGCCAG CAGCTGAGCT GGGTAGCCTC TCTGAGCTGT TTTCTGCCCC 2340
 CAAGGCTGG CTCTCCACAT CCAGGTGCAC CACTGAAGTG AGGACACACC GGAGCCAGGC 2400
 GCTGTCTAT GTTGAAGTGA GCTGTTTACA CCGCTCCGG AGAGCACCCC AGCGGCATCC 2460
 AGAAGCAGCT GCAAGTGTGC TGCCACCACC CTCTGCTCG CCTCTTCAA GTCTCTGTG 2520
 ACATTTTTTC TTTGCTCAGA AGCCAGGAAC TGGTGTCAAT CCTTAAAGA TACGTGCCGG 2580
 GGCCAGGTGT GGTGGCTCAC GCTGTAAAT CCAGCACTTT GGGAGGCCGA GGGGGCGGA 2640
 TCACAAAGT AGGACAGAGC CATCTGGCT AACACGGTGA AACCTGTCT CTACTAAAAA 2700
 TACAAAAAAA AATTAGCTAG GCGTAGTGGT TGGCACCTAT AGTCCAGCT ATCGGAAGG 2760
 CTGAAGCAGG AGAATGGTAT GAATCCAGGA GGTGGAGCT GCACTGAGCC GAGACCGTGC 2820
 CACTGCACTC CAGCCTGGGC AACACAGCGA GACTCCGTCT CSAGGAAAAA AAAAGAAAG 2880
 ACGGTACTT CCGGTGAGGA AGCTGGGCGC TGTTTTCGAG TTCAGGTGAA TTAGCCTCAA 2940
 TCCCGGTGT CACTGTCTCC CATAGCCCTC TTGATGGATC ACGTAAACT GAAAGGCAGC 3000
 GGGGAGCAGA CAAGATGAG GTCTACACTG TCCTTCAATG GGATTAAGC TATGGTTATA 3060
 TTAGCACCAA ACTTCTACAA ACCAAGCTCA GGGCCCCAAC CCTAGAGGG CCCAAATGAG 3120
 AGAATGGTAT TTAGGGATGG AAAACGGGCG CTGGCTAGAG CTTCGGGTGT GTGTGTCTGT 3180
 CTGTGTGTAT GCATACATAT GTGTGTATAT ATGGTTTTGT CAGGTGTGTA AATTGCAAA 3240
 TTGTTTCTCT TATATATGTA TGTATATATA TATATGAAA TATATATATA TATGAAAAAT 3300
 AAAGCTTAT TGTCCAGAAA AATCATACAT TGCTTTTTTA TTCTACATG GTACCAAGG 3360
 AACCTGGGG CCTGTGAAAC TACAACAAA AGGCACAAA AACCGTTCC AGTTGGCAGC 3420
 AGAGATCAGG GGTACTCTCT GCTTCTGAGC AATGGCTCA AGCTCTACCA GAGCAGACAG 3480
 CTACCTACT TTTACGACG AAAACGTCCT GTATGACGCA GCACGAAGG CCTGGCAGGC 3540
 TGTAGCAGG AGCTATGTCC CTTCCTATCG TTTCCGTCCA CTT

Seq ID NO: 120 Protein sequence:
 Protein Accession #: NP_006491.1

60
 65
 70
 75

1 11 21 31 41 51
 MGLPRLVCAF LLAACCCCPR VAGVPGEAEQ PAPELVEVEV GSTALLKCGL SQSQGNLSHV 60
 DWFSVHKEKR TLIFRVRRGQ GQSEPEGEYQ RLSLQDRGAT LALTQVTPQD ERIFLCQGRK 120
 PRSQEYRIQL RVYKAPBEPN IQVNPGLIPV NSKEPEEVAT CVGRNGYPIP QVIWYKNGRP 180
 LKEEKNRVHI QSSQTVESGG LYTLQSLIKA QLVKEDKDAQ FYCELNYRLP SGNHMKESRE 240
 VTVFVPFYTE KWLVEVPVG MLKEGDRVEI RCLADGNPPP HPSISKQNP S TREAEETTN 300
 DNGVLVLEPA RKEHSGRYEC QAWNLDTMIS LLSEPQELLV NYVSDVRVSP AAPERQEGSS 360
 LTLTCEABSS QDLFPWLRE ETDQVLERGP VLQLHDLKRE AGGGYRCVAS VPSIPGLNRT 420
 QLVKLAIFGP PWMAFKERKV WVKENMVLNL SCEASGHPRP TISWNVNGTA SEQDQDPQRV 480
 LSTIANVLTP ELLETGVECT ASNDLGRNTS ILFLVLVNL TLTPDSNTTT GLSTSTASPH 540
 TRANSTSTER KLPEPESRGV VIVAVIVCIL VLAVLGAVLY FLYKKGKLP RRSRQEIITL 600
 PPSRKTELVV EVKSDKLPEE MGLLQSSGD KRAPGDQGEK YIDLRLH

Seq ID NO: 121 DNA sequence
 Nucleic Acid Accession #: NM_018306
 Coding sequence: 60-671

80
 85

1 11 21 31 41 51
 ATAGTCTACA CAGAGCTCCC CTTGCTGCCC AGACAAGCTG AAGGACCACA GGAAAAGCCA 60
 TGGAGACTTC AGCATCTCTC TCCAGCCTC AGGACAACAG TCAAGTCCAC AGAGAAACAG 120
 AAGATGTAGA CTATGGAGAG ACAGATTTC ACAAGCAAGA CCGGAAGGCT GGACTCTTTT 180
 CCCAAGAA CAATATGAGAGA AACAACTCT CTTCCTCTCC CTCTCTTCC TCCTCATCTC 240
 CCTCATCTTC TCATCTCTC TCCTCTCTAG GTCTCGGCA TGGGAGCGCT GACGTTTTGA 300

5
10
15
20

```
AGGATGAGCT TCAACTCTAT GGAGATGCTC CTGGAGAGGT GGTACCCCTCT GGGGAATCAG 360
GACTCCGAG GAGAGGCTCT GACCCAGCAA GTGGAGAAGT GGAGGCTCT CAGTTAAGAA 420
GACTGAATAT AAAGAAAGAT GATGAGTTTT TCCATTTCGT CCTCTCTGTC TTTGCCATCG 480
GGGCTTGTCT GGTGTGTAT CACTATTACG CAGACTGGTT CATGTCTCTT GGGGTCTGCC 540
TGCTCACCTT CGCTCCCTCG GAAACCGTTG GCATCTACTT CGGACTAGTG TACCGTATCC 600
ACAGCGTCTT CCAAGGCTTC ATCCCCTCTT TCCAGAAATT TAGGCTGACA GGGTTCAGGA 660
AGACTGACTG AGGCCACTTC CAGGTGGGCA GCAGAGGCAG GCCCAGTGT GACCACCACT 720
GGCAGCCCTG AGCCCAAGG GGCAGAGCAG CATCTGAGA GACCCACAGG AGACCAAGCC 780
AGACCAATAA ACAGAACACT TTTCTTCCA TGTGGTCTGA ATGTTGGCAC CAGCCCGGSC 840
AGGGGCACTT CATTTGGGCA GTACTGCTGT GCAACCCAGC TGCAAGGATG GAAGGCAGAG 900
GGTGGGTGTG GGGCTGAGG CTTCACAGTA CCTGGACCAG CAGGAAGATT CTGGGAGTCT 960
ACTGCTCTCA GAGGACAGCA AGGACCCCTG AGCTCTGCAA GCTGTGATCT GTCTGGGTTT 1020
ATGTTTCTTC TCAATCCCA GGCTATCTGC ATGCGCTCTC AGGTGCTACC GAGCCATCCT 1080
GGGAGAGATG GATGTCCAC TGCTTTGAGG CAGGGAGCCA TCGGCTGGG GCCCTTGGT 1140
GAACTCTGAT CAGGTAAAGT GCTGAGGACT AAAACCATTT TTTTTCACCC CAAAAAATAA 1200
GGCAGGAAAA TGATCATCAG AAATAAATG GCAGCCAGGC ATGGGGGCTC ACAGCTGTAA 1260
TCTCGCACTT TTGGGAGGCT CAGGCTAAGG GTCGCTTGAA GCTGAGAGTT CAAGACCAAC 1320
CTGGGCAACA TAGTGAGACC CCCATCTCTA CAATTTTTTT TTAATGACCA AATGTGGCGG 1380
TACATACCTG TACATACCTG CGGTTCACGC TACTCAAGAG GCTGAGGCAG GAGGACTGCT 1440
TGAGCCAGG AGTTTCAAGG TGCAGTGAAG TACGATCAAG CCACTGCACT CCAGCTGGG 1500
CGACAGACGA AGATGTTTC TCTAAAAAT
```

Seq ID NO: 122 Protein sequence:
Protein Accession #: NP_060776

25
30

```
1 11 21 31 41 51
| | | | |
METSASSSQP QDNSQVHRET EDVDYGETDF HKQDGKAGLP SQEQYERNKS SSSSSSSSSS 60
SSSSSSSSSS GPGHGEPDVL KDEQLLYGDA PGEVVPSPGES GLRRRSGSDPA SGEVEASQLR 120
RLNIKKDDEF FHFVLLCPAI GALLVCYHYI ADWFMSLVGG LLTFASLETG GIYFLGVYRI 180
HSLVQGFIPL FQKRLTGFR KTD
```

Seq ID NO: 123 DNA sequence
Nucleic Acid Accession #: BC022542
Coding sequence: 243..896

35
40
45
50
55
60
65
70

```
1 11 21 31 41 51
| | | | |
ACTTGGTCCC AGCCGATAAA TCTGGGGCAG CGCGCGTAG GAGCTGCGGG CGGCCAGGCC 60
CCTTCTCTGG TCCGCACCTG GCGCCGCGCG CCCCTCTCGG GCGTCCGGCT TCCGGCGTCC 120
TGGCGGCTCG GGTGCGCGCG GTTCGGGCGG CGCGCTGGCT GCTCCTCGGG GCGGCGACGG 180
GGCTCACGCG CGGGCCCGCC ACGGCCTTCA CCGCCGCGCG CTCTGACGCC GGCATAAGGG 240
CCATGTGTTC TGAATATTAT TTGAGGCAAG AAGTTTGTAA AGATGGTTTC CACAGAGACC 300
TTTTAATCAA AGTGAAGTTT GGGGAAAGCA TTGAGGACTT GCACACGTGC CGTCTCTTAA 360
TTAAACAGGA CATTCCTGCA GGACTTTATG TGGATCCGTA TGAGTTGGCT TCATTACGAG 420
AGAGAAACAT AACAGAGGCA GTGATGGTTT CAGAAAATTT TGATATAGAG GCCCCTAACT 480
ATTTGTCCAA GGAGTCTGAA GTTCTCATTT ATGCCAGACG AGATTACAG TGCAATTGACT 540
GTTTTCAGCG CTTTTGGCCT GTGCACTGCC GCTATCATCG CGCGCACAGT GAAGATGGAG 600
AAGCCTCGAT TGTGTCTCAAT AACCCAGATT TGTGTATGTT TTGTGACCAA GAGTTCCCGA 660
TTTTGAAATG CTGGGCTCAC TCAGAAAGTG CAGCCCGCTG TGCTTTGGAT AATGAGGATA 720
TATGCCAATG GAACAGAGT AAGTATAAAT CAGTATATAA GAATGTGATT CTACAGATT 780
CAGTGGGACT GACTGTACAT ACCTCTCTAG TATGTTCTGT GACTCTGCTC ATTACAATCC 840
TGTGCTCTAC ATTGATCCTT GTAGCAGTTT TCAAAATATG CCATTTTTC CTATAAGTTT 900
TATGTAGTTA AATGCTTCTT AGAAACCTAA ATAAGATCTA TTAATTTCTG ACGAGAGGTG 960
TTCTCTAGA ATTAATTAAT TTTATCTTTT GTCTTCATTT GTGGCCAAAA TTATGTTTAC 1020
TAGAGGAATG TTGGATGACT TCTCAGCTAA TTCCAAAATG TAGTGCTCTA TTGCATGGAT 1080
CCTTGGTAAT CCTCAAGCAT CAGATGCCAT AAGGGGAAAC TTAATTTCTG TAAATTAATG 1140
TTTATTTTGT GAGAAGTGAC TTTATCTTCA TTTGGGGTAG AAAAATTTAT TCTTTATGTA 1200
GTAGAGACAA ATTATTTCTA TTTTGCAAGT ACTTTCAATT TAAGCTACAA ATTGAGAAAA 1260
CGGTATATAA TAAGAAATAA ATAGGCCAGG CACAGTGGCT CACACCTGTA ATCCAGCAC 1320
TTTGGGAGGC CGAGGTGGGC GGATCACCAG AGGTCAAGAG TTTGAGACCA GCTTGGTGAA 1380
ACCTGTCTCT TACTAAAAAT ACAAAGTTA GCTGGGGCTG GTGGTGGGCA TCTGTAGTCC 1440
CAGCTAATTG GAAGGGTGAG GCGGGAGGAT CGCTTGAACC TGGGAGGCG AGGTTCACGA 1500
GAGCCAAGAT CGCACCACCT CACTACAGCC TGGGCGACAG AACGAGACCC TGTCTCCAAA 1560
GGAAAAACAA AAAAGAAAG TAAATAAAT TGGATGAAAA TCATGTTTAT TTAATAGTA 1620
ATGTCATGAG ACTATTAAAG ATGTGCCAGA GTTCAATGA AAATCATTAA AGTAGGACAG 1680
CTAAGAAATT AATATTAATA TAAAAATTAT TGATAATCTT AAATATTGA TTATTCCTTA 1740
ACGCATCCA TTCTCTTTT ACATTTTATC ATGTTTCTT TGAATATATG AATTGGCAAA 1800
GGACTTGATG AAATCGAGTA CTAAGATTG GTACAGAGTA TGTGAGGAG ACAACTCAGA 1860
TTGCCATTTT AAATAAGTT GTACATGAAC AAAAAAATAA AAAAAA
```

Seq ID NO: 124 Protein sequence:
Protein Accession #: AAH22542

75
80

```
1 11 21 31 41 51
| | | | |
MCSEILRQE VLKDGPHRDL LIKVKFGESI EDLHTRLLI QDIPAGLYV DPYELASLR 60
RNITEAVMVS ENFDIEAPNY LSKSEVLIY ARRDSQCIDC FQAPLPVHCR YHRPHSEDE 120
ASIVVMNPD LMFCDQAGSR MIRFRFDSF DKTIEFPILK CWAHSEVAAP CALENEDICQ 180
WKKMKYSKY KVVILQVPGV LTVHTSLVCS VTLLITILCS KKKKK
```

Seq ID NO: 125 DNA sequence
Nucleic Acid Accession #: NM_004994.1
Coding sequence: 20..2143

85

1 11 21 31 41 51
AGACACCTCT GCCCTCACC TGAGCCTCTG GCAGCCCTGT GTCTGGTGCT TCCTGGTGCT 60
5 GGGCTGCTCT TTTGCTGCCC CCAGACAGCG CCAGTCCACC CTGTGTCTCT TCCTGGGAGA 120
CCTGAGAACCC AATCTCACCG ACAGGCAGCT GGCAGAGGAA TACCTGTACC GCTATGGTTA 180
CACTCGGGTG GCAGAGATGC GTGGAGAGTC GAAATCTCTG GGGCCTGCGC TGCTGCTCTC 240
CCAGAAGCAA CTGTCCCTGC CCGAGACCGG TGAGCTGGAT AGCCGCCACG TGAAGGCCAT 300
GCGAACCCCA CGGTGGGGGG TCCAGACCTT GGCAGATTG CAAACCTTTG AGGGCGACCT 360
10 CAAGTGGCAC CACCACAACA TCACCTATTG GATCCAAAAC TACTCGGAAG ACTTCCCGGG 420
GGCGGTGATT GAGGAGCGCT TTGCCCGCGC CTTCGCACTG TGGAGCGGGG TGACGCGCGT 480
CACCTTCACT CGGTGTACAC GCGCGGACGC AGACATCGTC ATCCAGTTTG GTGTGCGGGA 540
GCACGAGAGC GGGTATCCCT TCGACGGGAA GGACGGGCTC CTGGCACACG CCTTCTCTCC 600
TGGCCCGCGC ATTACGGGAG ACGCCCATTT CGACGATGAC GAGTTGTGGT CCCTGGGCAA 660
15 GGGCGTCTGT GTTCCAATCT GGTTTGAAA CGCAGATGGC GCGGCGTCC ACTTCCCTTT 720
CATCTTCGAG GGGCGCTCCT ACTCTGCTG CACCACCGAC GGTGCTCCG ACGGCTTGCC 780
CTGGTGCAGT ACCACGGCCA ACTACGACAC CGACGACCGG TTTGGTCTCT GCCCCAGCGA 840
GAGACTCTAC ACCCGGGAGC CCAATGCTGA TGGGAAACCC TGCCAGTTTC CATTCATCTT 900
CCAAGGCCAA TCCTACTCGC CCTGCACCA GGCAGGTGCG TCCGACGGCT ACCGCTGGTG 960
20 CGCCACCACC GCCAACTACG ACCGGGACAA GCTCTTCGCG TCTTCCCGA CCGAGCTGA 1020
CTCGACGGTG ATGGGGGGCA ACTCGGGGGG GGAGCTGTGC GTCTTCCCCT TCACCTTCCCT 1080
GGTAAGGAGT TACTCGACCT GTACCAAGCA GGGCCGCGGA GATGGGCGCC TCTGGTGCGC 1140
TACCACCTCG AACTTTGACC GCGACAAGAA GTGGGGCTTC TGCCCGGACC AAGGATACAG 1200
TTTGTTCCTC GTGGCGGCGC ATGAGTTGCG CCACGCGCTG GGTATGATC ATTCTCTAGT 1260
25 GCGGAGGGCG CTCATGTACC CTATGTACCG CTTCAGTGA GGGCCCCCTT TGCATAAGGA 1320
CGACGTGAAT GGCATCGGCG ACCTCTATGG TCCTCGCCCT GAACCTGAGC CACGGCTTCC 1380
AACCACCACC ACACCGCAGC CCACGGCTCC CCGGACGGTC TGCCCCACCG GACCCCCAC 1440
TGTCCACCCC TCAGAGCGGC CCACAGCTGG CCCCACAGGT CCCCCCTCAG CTGGCCCCAC 1500
AGGTCCCCC ACTGCTGGCC CTCTACGGC CACTACTGTG CCTTTAGTC CGGTGGAAGA 1560
30 TGCGTGAAC GTGAACATCT TCGACGCCAT CGCGAGATT GGAACACAGC TGTATTGTGT 1620
CAAGGATGGG AAGTACTGGC GATTCTCTGA GGGCAGGGG AGCGGGCGCC AGGGCCCCCT 1680
CCTTATCGCC GACAAGTGGC CCGCGCTGCC CCGCAAGCTG GACTCGGTCT TTGAGGAGCC 1740
GCTCTCAAG AAGCTTTTCT TCTTCTCTGG GCGCAGGTG TGGGTGTACA CAGGCGCGTC 1800
GGTCTGGGC CCGAGGCGTC TGGACAAGCT GGGCCTGGGA GCGGACGTGG CCGAGGTGAC 1860
35 CGGGGCCCCC CGGAGTGGA GGGGGAAGAT GCTGCTGTT AGCGGGCGGC GCCTCTGGAG 1920
GTTGACGCTG AAGGCGCAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGGATGTT 1980
CCCCGGGGTG CCTTTGGACA CGCAACGAGT CTTCAGTAC CGAGAGAAAG CCTATTCTG 2040
CCAGGACCGC TTCTACTGGC GCGTGAATTC CCGGAGTGAG TTGAACACAG TGGACCAAGT 2100
GGGCTACGTG ACCTATGACA TCCTGCAGTG CCTTGAGGAC TAGGGCTCCC GTCTGCTTT 2160
40 GCAGTGCAT GTAAATCCCC ACTGGGACCA ACCCTGGGGA AGGAGCCAGT TTGCGGATA 2220
CAAACTGGTA TTCTGTTCTG GAGGAAAGGG AGGAGTGGAG GTGGGCTGG CCCTCTCTTC 2280
TCACCTTTGT TTTTGTGTGG AGTGTCTTA ATAACTTGG ATTCTTAAC CTTT

Seq ID NO: 126 Protein sequence:
Protein Accession #: NP_004985.1

1 11 21 31 41 51
MSLWQPLVLV LLVLGCCFAA PRQRQSTLVL FPGDLRLNLT DRQLAEELY RYGYTRVAEM 60
50 RGEKSLGPA LLLQKQLSL PETGELDSAT LKAMRTFRCG VPDLRFTQF EGDLLKHHEN 120
ITYWIQNYSE DLPRAVIDDA FARAPALWSA VTLPTFRVY SRDADIVIQF GVAEHGEGYP 180
FDGKDLGAH AFPPGPGIQG DAHFDDDELW SLGKGVVPT RFGNADGAAC HPFFIFEGRS 240
YSACTDGRS DGLPWCSTTA NYDTRDRFGF CPSERLYTRD GNADGKPCQF PFIQQQSYS 300
ACTTDRSDG YRWCAATTANY DRDKLFGFCP TRADSTVMGG NSAGELCVFP FTFLGKEYST 360
55 CTSEGRGDR LWCATSNFED SDKKWGFPCD QGYSLELVAA HEFHALGLD HSSVPEALMY 420
PMYRFTGPP LHKDDVNGIR HLYGPRPEPE PRPPTTTTPQ PTAPPTVCP TGPPTVHPSER 480
PTAGPTGPPS AGPTGPPTAG PSTATTVPLS PVDDACNVNI FDAIAEIGNQ LYLFDGKYW 540
RFSEGRGSRP QGFPLIADKW PALPRKLDV FEEPLSKLFP FFSGRQVWVY TGASVLGPRR 600
60 LDKLGLGADV AQVTGALRSR RGKMLLFSGR RLWRFDVKAQ MVDPRSASEV DRMPFVPLD 660
THDVQYREK AYFCQDRPYW RVSSRSELNQ VDQVGYVYTD ILQCPED

Seq ID NO: 127 DNA sequence
Nucleic Acid Accession #: NM_004181
Coding sequence: 32-670

1 11 21 31 41 51
GCAGAAATAG CTAGGGGAGA TCAACCCCGA GATGCTGAAC AAAGTGCTGT CCCGGCTGGG 60
65 GGTGCGCGCG CAGTGGCGCT TCGTGGACGT GCTGGGGCTG GAAGAGGAGT CTCTGGGCTC 120
GGTGCCACGG CCTGCGCTGC CGCTGCTGCT GCTGTTTCCC CTACGGGCC AGCATGAGAA 180
70 CTTAGGAAA AAGCAGATTG AAGAGCTGAA GGCACAAGAA GTTAGTCTTA AAGTGTACTT 240
CATGAAGCAG ACCATTGGGA ATTCTGTGG CACAATCGGA CTTATTACG CAGTGGCCAA 300
TAATCAAGAC AAATGGGGAT TTGAGGATGG ATCAGTTCTG AAACAGTTTC TTTCTGAAAC 360
AGAGAAAATG TCCCTGAAG ACAGAGCAA ATGCTTTGAA AAGAATGAGG CCATACAGGC 420
75 AGCCCATGAT GCGGTGGCAC AGGAAGGCCA ATGTCGGGTA GATGACAAGG TGAATTTCCT 480
TTTTATTCTG TTTAACAACG TGGATGGCCA CCTCTATGAA CTGTATGAGC GAATGCCCTT 540
TCCGTTGAAC CATGGCGCCA GTTCAGAGGA CACCTGCTG AAGGACGCTG CCAAGGTGTG 600
CAGAGAATTC ACCGAGCGTG AGCAAGGAGA AGTCCGCTTC TCTGCGGTGG CTCTCTGCAA 660
GGCAGCTTAA TGCTCTGTGG GAGGGACTTT GCTGATTTC CCTCTTCCCT TCAACATGAA 720
AATATATACC CCCCATGCAG TCTAAATGCT TTCAGTACT GTGAAACACA GCTGTTCTTC 780
80 TGTCTGAGC ACACGCTTTC CCTCAGCCA CACCCAGGCA CTTAAGCACA AGCAGAGTGC 840
ACAGCTGTCC ACTGGGCCAT TGTGGTGTGA GCTTCAGATG GTGAAGCATT CTCGCCAGTG 900
TATGCTGTGT ATCCGATATC TAACGCTTTA AATGGCTACT TTGGTTTCTG TCTGTAAGTT 960
AAGACCTTGG ATGTGGTTAT GTTGTCTTAA AGAATAAATT TTGCTGATAG TAGC

Seq ID NO: 128 Protein sequence:
Protein Accession #: NP_004172

1 11 21 31 41 51
MLNKVLSRLG VAGQWRFVDV LGLEESLSGS VPAPACALLL LPPLTAQHEN FRKKQIEELK 60
GQEVSPKVYP MKQTIGNSCG TIGLIHAVAN NQDKLGFEDG SVLKQFLSET EKMSPEDRAK 120
CFENEAIOA AHDVAQEGQ CRVDDKVNPH FILFNNVDGH LYELDGRMPF PVNHGASSED 180
TLKDAARVC REPTEREQGS VRFAVALCK AA

Seq ID NO: 129 DNA sequence
Nucleic Acid Accession #: NM_000213
Coding sequence: 127-5385

1 11 21 31 41 51
CGCCCGCGCG CTGCAGCCCG ATCTCCTAGC GGCAGCCCG GCGCGGAGGG AGCGAGTCCG 60
CCCGAGGGTA GGTCCAGGAC GGGGCGACAG CAGCAGCCGA GGCTGGCGGG GAGAGGGAGG 120
AAGAGGATGG CAGGGCCACG CCCAGCCCA TGGGCCAGGC TGCTCTGCGC AGCCTTGATC 180
AGCGTCAGCG TCTCTGGGAC CTCTGGCAAC CGCTGCAGAA AGGCCCCAGT GAAGAGCTGC 240
ACGGAGTGTG TCGGTGTGGA TAAGGACTGC GCCTACTGCA CAGACGAGAT GTTCAGGGAC 300
CGCGCTGCA ACACCCAGCG GGAGCTGCTG GCGCGGGGCT GCCAGCGGGA GAGCATCGTG 360
GTCACTGAGA CGAGCTTCCA AATCACAGAG GAGACCCAGA TTGACCCAC CCTGGCGGCG 420
AGCCAGATGT CCCCCAAGG CCTGCGGGTC CGTCTGCGGC CCGGTGAGGA GCGGCATTTT 480
GAGCTGGAGG TGTGTGAGCC ACTGGAGAGC CCCGTGGACC TGTACATCCT CATGGACTTC 540
TCCAACCTCA TGTCCGATGA TCTGGACAAC CTCAAGAAGA TGGGGCAGAA CCTGGCTCGG 600
GTCTGAGCC AGCTACCCAG GACTACACT ATTGGATTGG GCAAGTTTGT GGACAAAGTC 660
AGCGTCCCCG AGACCGACAT GAGGCTGAG AAGCTGAAG AGCCCTGGCC CAACAGTGAC 720
CCCCCTTCT CTTTCAAGAA CGTCATCAGC CTGACAGAAG ATGTGGATGA GTTCCGGAAT 780
AAACTCAGG GAGAGCGGAT CTCAGGCAAC CTGGATGCTC CTGAGGGCGG CTTTGATGCC 840
ATCTGCGAGA CAGCTGTGTG CAAGAGGGAC ATTGGCTGGC GCGCGGACAG CACCCACTTG 900
CTGGTCTTCT CCACCGAGTG AGCCTTCCAC TATGAGGCTG ATGGCGCCAA CGTGTGGGCT 960
GGCATCATGA GCGCGCAACG TGAACGGTGC CACTGGACA CCACGGGCAC CTACACCCAG 1020
TACAGGACAC AGGACTACCC GTGGGTGCCC ACCCTGGTGC GCCTGCTGCG CAAGCACAAC 1080
ATCATCCCCA TCTTTCGTGT CACCAACTAC TCCTATAGCT ACTACGAGAA GCTTCACACC 1140
TATTTCCCTG TCTCTCACT GGGGGTGTCT CAGGAGGACT CGTCCAACT CGTGGAGCTG 1200
CTGAGGAGG CCTTCAATCG GATCCGCTCC AACCTGGACA TCGGGGCCCT AGACAGCCCC 1260
CGAGGCTTTC GACAGAGGT CACTCCAAAG ATGTTCCAGA AGACGAGGAC TGGGTCTTTT 1320
CACATCCGCG GGGGGGAAAT GGGTATATAC CAGGTGCAGC TGGCGGCCCT TGAGCACGTG 1380
GATGGGAGCG ACGTGTGCCA GCTGCCGAGG GACCAAGAGG GCAACATCCA TCTGAAACCT 1440
TCTTCTCCCG ACGGCTCAA GATGGAAGCG GGCATCATCT GTGATGTGTG CACTTGCGAG 1500
CTGCAAAAG AGGTGCGGTG AGCTCGCTGC AGCTTCAACG GAGACTTCGT GTGCGGACAG 1560
TGTGTGTGCA GCGAGGCTGC GAGTGGCCAG ACCTGCAACT GCTCCACCG CTCTCTGAGT 1620
GACATTACAG CCTGCTGCG GAGGGGCGAG GACAAGCCGT GCTCCGCGCG TGGGGAGTGC 1680
CAGTGGCGGG ACTGTGTGTG CTACGGCGAA GGGCGCTAGC AGGGTCAGTT CTGCGAGTAT 1740
GACAACTTCC AGTGTCCCG CACTTCCGGG TTCTCTGCA ATGACCGAGG ACGCTGCTCC 1800
ATGGGCGAGT GTGTGTGTGA GCCTGGTGGG ACAGGCCCCA GCTGTGACTG TCCCTCTCAG 1860
AATGCCACCT GCATCGACAG CAATGGGGCG ATCTGTAATG GAGCTGGCCA CTGTGAGTGT 1920
GGCGCTGCC ACTGCCACCA GCAGTGGCTC TACACGGACA CCATCTGCGA GATCAACTAC 1980
TGGCGATCC ACCCGGGGCT CTGCGAGGAC CTACGCTCCT GCGTGCAGTG CCAGGCGTGG 2040
GGCAGCGCGG AGAAGAAGGG GCGCACTGTG GAGGAATGCA ACTTCAAGGT CAAGATGGTG 2100
GAGAGCTTAA AGAGAGCGCA GGAGGTGGTG GTGCGCTGCT CCTTCCGGGA CGAGGATGAC 2160
GACTGCACCT ACAGCTCAC CATGGAAGGT GACGGCGCCC CTGGGCCCAA CAGCACTGTC 2220
CTGGTGACA AGAAGAAGGA CTGCCCCCG GGTCTCTCT GGTGGCTCAT CCCCCTGCTC 2280
CTCTCTCTCC TGCGCTCCT GGGCTGTGTA CTGCTGTAT GCTGGAAGTA CTGTGCTGCG 2340
TGCAAGGCTT GCTGGGCTCT TCTCCGCTGC TGCAACCGAG GTCACTAGGT GGGCTTTAAG 2400
GAAGACCTCT ACATGCTCG GGAAGAACCTG ATGGCTCTG ACCACTTGA CAGCGCCATG 2460
CTGCGCAGCG GGAACCTCAA GGGCGTGC GTGGTCCGCT GGAAGTCA CACACATG 2520
CAGCGCGCTG GCTTTGCCAC TCATGCCGCC AGCATCAACC CCACAGAGCT GGTGCCCTAC 2580
GGGCTGTCTT TGCGCTTGGC CGGCTTTGC ACCGAGAAC TGCTGAAGCC TGACACTGG 2640
GAGTGGCGCC AGCTGCGGCA GGAGGTGGAG GAGAACCTGA ACGAGGTCTA CAGGCAGATC 2700
TCCGTGTGAC ACAAGCTCCA GCAGACCAAG TTCGGCAGC AGCCCAATGC CGGGAJAAAG 2760
CAAGACCA CAATGTGGA CACAGTGTG ATGGCGCCCC GCTCGGCCAA GCGCGCCCTG 2820
CTGAAGCTTA CAGAGAAGCA GGTGGAACAG AGGGCTTCC ACAGCTCAA GGTGGCCCC 2880
GGCTACTACA CCTCACTGC AGACAGGAC GCGCGGGCA TGGTGGAGTT CCAGGAGGGC 2940
GTGGAGCTGG TGGAGTACG GGTGCCCTC TTTATCCGGC CTGAGGATGA CGACGAGAAG 3000
CAGCTGTGCG TGGAGGCCAT CGACGTGCC GCAGGCACTG CCACCTCGG CCGCGGCTG 3060
GTAAACATCA CCATCATCAA GGAGCAAGCC AGAGAGCTGG TGTCTTTGA CGAGCTGAG 3120
TTCTGGTCA GCGCGGGGA CCAGGTGGCC CGCATCCCTG TCATCCGCGG TGTCTGGAC 3180
GGCGGGAAGT CCCAGGTCTC CTACCGACA CAGGATGGCA CCGCGCAGG CAACCGGGAC 3240
TACATCCCCG TGGAGGGTGA GCTGCTGTT CAGCTGGGG AGGCTTGAA AGAGCTGCAG 3300
GTGAAGCTCC TGGAGTGA AGAAGTTGAC TCCCTCTGCG GGGCGCGCA GGTCCGCGT 3360
TTCCAGCTCC AGCTCAGCAA CCTAAGTTT GGGGCCACC TGGGCCAGCC CCACTCCACC 3420
ACCATCATCA TCAGGAGACC AGATGAAGT GACCGGAGCT TCAAGATCA GATGTGTGCA 3480
TCACAGCCAC CCCCTCAGG CGACCTGGGC GCGCGCAGA ACCCAATGC TAAGGCGGCT 3540
GGGTCCAGGA AGATCCATT CAACTGGCTG CCCCCTTCTG GCAAGCCAAT GGGGTACAGG 3600
GTAAAGTACT GGATTCAGGG TGAATCGGAA TCCGAAGCCC ACCTGCTCGA CAGCAAGGTG 3660
CCCTCAGTGG AGCTCACCAA CTTGTACCG TATTGCGACT ATGAGATGAA GGTGTGGGCC 3720
TACGGGGCTC AGGGGAGGG ACCCTACAGC TCCCTGGTGT CCGCGCGAC CCACCGAGAA 3780
GTGCCAGCG AGCCAGGGCG TCTGGCTTTC AATGTGCTCT CCTCCAGGT GACCCAGCTG 3840
AGCTGGGCTG TGGCGGCTGA GACCAACGCT GAGATCAAG CTTACAGGT CTGCTATGGC 3900
CTGGTCAACG ATGACAACCG ACCTATTGGG CCAATGAAGA AAGTGTGGT TGACAACCT 3960
AAGAACCAGA TGCTGCTTAT TGAGAACCTT CGGAGTCCC AGCCCTACG CTACAGGTG 4020
AAGGGCGCA ACGGGGCGG CTGGGGGCTG GAGCGGGAGG CCATCATCAA CCGGCCACC 4080
CAGCCCAAGA GCGCCATGTC CATCCCCATC ATCCCTGACA TCCCTATGTT GGACGCCAG 4140
AGCGGGGAGG ACTACGACAG CTTTCTTATG TACAGCGATG ACGTTCTACG CTCTCCATCG 4200
GGCAGCCAGA GCGCCAGCGT CTCCGATGAC ACTGAGCACC TGGTGAATGG CCGGATGAG 4260
TTTGCTTCC CCGGCGAGC CAACTCCCTG CACAGGATGA CCAAGACAG TGCTGCTGCC 4320
TATGCAACCC ACCTGAGGCC ACACGTGCC CACCGGTGTC TAAGCACATC CTCCACCTC 4380

WO 02/086443

PCT/US02/12476

5
10
15
20
25

ACACGGGACT ACAACTCACT GACCGGCTCA GAACACTCAC ACTGACACAC ACTGCGGAGG 4440
GACTACTCCA CCTTCACTCT CGTCTCTCTCC CACGACTCTC GCTGACTGTC TGGTGTGCCC 4500
GACACGGCCA CCGGCTTGGT GTTCTCTGTC CTGGGGGCCA CATCTCTCAG AGTGAGCTGG 4560
CAGGAGCGCG GGTGCGAGCG GCCGCTGCAG GGCTACAGTG TGGAGTACCA GCTGCTGAAC 4620
GGCGGTGAGC TGCATCGGCT CAACATCCCC AACCTGCCC AGACCTCGGT GGTGGTGGAA 4680
GACCTCTCTG CCAACCACTC CTACGTGTTC CGGTGCGGG CCCAGAGCCA GGAAGGCTGG 4740
GGCGGAGAGC GTGAGGCTGT CATCACCATT GAATCCAGG TGCACCCGCA GAGCCCACTG 4800
TGTCCCTGTC CAGGCTCGC CTTCACCTTG AGCACTCCCA GTGCCCCAGG CCCGCTGCTG 4860
TTCACCTCCC TGAGCCGAGA CTCGCTGCAG CTGAGCTGGG AGCGGCCACG GAGGCCCAAT 4920
GGGATATCG TCGGCTACCT GGTGACCTGT GAGATGGCCC AAGGAGGAGG GCCAGCCACC 4980
GCATTCCGGG TGGATGGAGA CAGCCCGAG AGCGGCTGA CGGTGCGGG CCTCAGCGAG 5040
AACGTGCCCT ACAAGTTCAA GGTGCAGGCC AGGACCACTG AGGGCTTCGG GCCAGAGCGC 5100
GAGGGCATAGA TCACCATAGA GTCCGAGGAT GGAGGACCTT TCCCGCAGCT GGGCAGCCGT 5160
GCCGGGCTCT TCCAGCACCC GCTGCAAAAG GAGTACAGCA GCATCACAC CACCACACCC 5220
AGCGCCACCG AGCGCTTCTT AGTGGATGGG CCGACCTGG GGGCCAGCA CCTGAGGCA 5280
GGCGGCTCCC TCACCCGGCA TGTGACCCAG GAGTTTGTGA GCGGACACT GACCACACCC 5340
GGAACCTTA GCACCCATAT GGACCAACAG TTCCTCCAAA CTTGACCGCA CCTGCCCA 5400
CCCCGCCAT GTCCCACTAG GCGTCTCTCC GACTCTCTC CGGAGGCTC CTCAGCTACT 5460
CCATCCTTGC ACCCTGGGG GCCCAGCCCA CCGCATGCA CAGAGCAGG GCTAGGTGTC 5520
TCCTGGGAGG CATGAGGGG GCAAGGTCCG TCCTCTGTGG GCCCAAACTT ATTTGTAAAC 5580
AAAGAGCTGG GAGCAGCACA AGGACCCAGC CTTTGTCTG CACTTAATAA ATGGTTTTCG 5640
ACTG

Seq ID NO: 130 Protein sequence:
Protein Accession #: NP_000204

30
35
40
45
50
55
60

1 11 21 31 41 51
MAGPRPSPWA RLLLAALISV SLSGTLANRC KKAAPVKSCTE CVRVKDCAY CTDEMFRDRR 60
QNTQAEELAA GCQRESIVVM ESSFOITEET QIDTTLRRSQ MSPQGLRVRL RPEGERHFEL 120
EVFEPLESPV DLYILMDFSN SMSDDLNLK KMQQNLARVL SOLTSDYTTIG FGKFPVKVSV 180
PQTDMPREKL KPWPNSDPP FSKNVISLT EDVDEPRNKL QGERISGNLD APEGGFDAIL 240
QTAVCTRDIG WRPDSTHLIV FSTESAFHYE ADGANVLAGI MSRNDERCHL DTTGTYYTQR 300
TQDVPSPVPL VRLAKHNI PIFAVTNYSY SYYEKLHYE FVSSLGVLEQ DSSNIVELLE 360
EAFNRIRSNL DIRALDSPRG LRTEVTSKMF QKTRTGSFHI RRGEVGIYQV QLRALHVDG 420
THVQQLPEDQ KGNHILKPSF SDGLKMDAGI ICDVCTCELQ KEVRSARCSF NGDFVCGQCV 480
CSEGSWGQTC NCSTGSLSDI QPCLREGEDK PCSGRGECQC GHCVCYGEGR YEGQFCEYDN 540
FQCPRTSGFL CNDGRGRCSMG QCVCEPGWTG PSCDCPLSNA TCIDSNGGIC NGRGHCECGR 600
CHCHQQLSYL DTICEINYSY IHPGLCEDLR SCVQCQAWGT GEKKGRTECE CNFKVMVDE 660
LKRAEEVVVR CSFRDEDDC TYSYMEGDG APGPNSTVLV HKKKDCPPGS FWLILPLLLL 720
LLPLLALLLL LCWKYCACCK ACLALLPCCN RGHMVGFEDL HYMLRENLMH SDHLDTPMLR 780
SGNLKGRDVV RWKVTNNMQR PGFATHAASI NPTLVVYGL SLRLARLCTE NLLKPDTRER 840
AQLRQEVEEN LNEVYRQISG VHKLQQTFR QQPNAKQKD BTIVDTVLMA PRSAKPAALK 900
LTERQVEQRA FHDLVAPGY YTLTADQDAR GMVEFQEGVE LVDVRVPLFI RPEDDDKQL 960
LVEAIDVPAG TATGLRRLVN ITI IKEQARD VVSFEQPEFS VSRGDQVARI PVIRVRLDGG 1020
KSQVSRYTQD TQAQGNRYI PVBGELLFQP GEANKELQVK LLELQEVDSL LRGRQVRRFH 1080
VQLSNPKFGA HLGQPHSTTI IIRDPDELDR SFTSQMLSSQ PPPHGLDLAG QNPNAKAAGS 1140
RKIHFNWLPF SGKPMGRYVK YWQGDSESE AHLDSKVPVS VELTNLYPYC DYEMKVCAYG 1200
AQGEGPYSSL VSCRTHQVVP SEPGRLAFNV VSSTVTQLSW AEPATNGEI TAYEVCYGLV 1260
NDNDRPIGPM KKVIVNDPNK RMLLIENLRE SQPYRYTVKA RAGAGWGPER EAIINLATOP 1320
KRPMSPPIIP DIPIVDAQSG EDYDSPLMYS DDVLRSPSGS QRPSVSDDTL HLVNRMDFN 1380
FPGSTNSLHR MTTSAAYG THLSPHVPHR VLSTSTLTR DYNLSTRSEH SHSTTLPRDY 1440
STLTSVSSHD SRLTAGVPDT PTRLVFSALG PTLRVSQWE PRCPRLQGY SVEYQLLNGG 1500
ELKRLNIPNP PSAPGLVET ALSPDSLQLS WERPRRENGD IVGYLVTCME AQGGGPATAF 1560
LFGSAFTLST PSAPGLVET ALSPDSLQLS WERPRRENGD IVGYLVTCME AQGGGPATAF 1620
RVDGDSPESE LTVPGLESENV PYKFKVQART TEGFGPEREG IITIESQDGG PFPQLGSRAG 1680
LFQHPLOSEY SSITTTHTSA TEPFLVDGPT LGAQHLEAGG SLTRHVTQEF VSRTLTTSST 1740
LSTHMDQPPF QT

Seq ID NO: 131 DNA sequence
Nucleic Acid Accession #: BC004372
Coding sequence: 132..2231

65
70
75
80
85

1 11 21 31 41 51
CCTCGTGCCG CGGACCCAG CCTCTGCCAG GTTCGGTCCG CCATCCTCGT CCGCTCCTCC 60
GCCGGCCCCC GCCCGCGGCC CAGGGATCCT CCAGCTCCTT TCGCCCGCGC CCTCCGTTGG 120
CTCCGACAC CATGGACAAG TTTTGTGGC ACGCAGCCTG GGGACTCTGC CTGTCGCCGC 180
TGAGCCTGGC GCAGATCGAT TTGAATATAA CCGCGCGCTT TGCAGGTGTA TTCCAGCTGG 240
AGAAAAATGG TCGCTACAGC ATCTCTCGGA CGGAGGCCGC TGACCTCTGC AAGGCTTTC 300
ATAGCACCTT GCCCAACAAT GCCCAGATGG AGAAGCTCT GAGCATCGGA TTTGAGACCA 360
GCAGGTATGG GTTCATAGAA GGGCATGTGG TGATTCCCG GATCCACCCC AACTCCATCT 420
GTGCAGCAAA CAACACAGGG GTGTACATCC TCACATCCAA CACTCCCAG TATGACACAT 480
ATTGCTTCAA TGTCTCAGCT CCACCTGAAG AAGATTGTAC ATCAGTCACA GACCTGCCA 540
ATGCCCTTGA TGGACCAATT ACCATAACTA TTGTTAACCG TGATGGCACC CGCTATGTCC 600
AGAAAGGAGA ATACAGAAGC AATCCTGAAG ACATCTACCC CAGCAACCTT ACTGATGATG 660
ACGTGAGCAG CGGCTCCTCC AGTGAAAGGA GCAGCACTTC AGGAGGTAC ATCTTTTACA 720
CCTTTTCTAC TGTACACCCC ATCCAGACG AAGACAGTCC CTGGATCACC GACAGCACAG 780
ACAGAAATCC TGTACCAAT AGCTTCTCAA ATACCATCTC AGCAGGCTGG GAGCCAAATG 840
AAGAAATGA AGATGAAAGA GACAGACACC TCACTTTTTC TGATCAGGC ATTGATGATG 900
ATGAAGATT TATCTCCAGC ACCATTTCAA CCACACACG GGCCTTTGAC CACACAAAAC 960
AGAACCAGGA GTGACACCCG TGAACCCCAA GCCATTCAA TCGGAAGTGT CTACTTCAGA 1020
CAACCACAG GATGACCTAT GTAGACAGAA ATGGCACCAC TGCTTATGAA GGAACCTGGA 1080
ACCCAGAACG ACACCTCTCC CTCAATCACC ATGAGCATCA TGAGGAAGAA GAGACCCAC 1140
ATTCTACAAG CACAATCCAG GCAACTCTTA GTAGTACAAC GGAAGAAACA GCTACCCAGA 1200
AGGAACAGTG GTTTGGCAAC AGATGGCATG AGGATATCG CCAACACCC AGAGAAGACT 1260

WO 02/086443

PCT/US02/12476

5
10
15
20

```

CCCATTCGAC AACAGGAGCA GCTGCAGCCT CAGCTCATAC CAGCCATCCA ATGCAAGGAA 1320
GGACAAACACC AAGCCAGAG GACAGTTTCT GGACTGATT CTTCACCCA ATCTCACACC 1380
CCATGGGACG AGGTGCATCA GCAGGAAGAA GGATGGATAT GGACTCCAGT CATAGTACAA 1440
CGCTTCAGCC TACTGCAAAAT CCAAAACACAG GTTTGGTGA AGATTGGAG AGGACAGGAC 1500
CTCTTTCAAT GACAACGCGAG CAGAGTAATT CTCAGAGCTT CTCTACATCA CATGAAGGCT 1560
TGAAGAAGA TAAAGACCAT CCAACAATT CTACTCTGAC ATCAAGCAAT AGGAATGATG 1620
TCACAGGTGG AAGAAGAGAG CCAAAATCATT CTGAAGGCTC AACTACTTTA CTGGAAGGTT 1680
ATACCTCTCA TTACCCACAC ACGAAGGAAA GCAGGACCTT CATCCCACTG ACCTCAGCTA 1740
AGACTGGGTC CTTTGGAGTT ACTGCAGTTA CTGTTGGAGA TTCCAACCTT AATGTCAATC 1800
GTTCTTATTC AGGAGACCAA GACACATTCC ACCCCAGTGG GGGTCCCAT ACCACTCATG 1860
GATCTGAATC AGATGGACAC TCACATGGGA GTCAAGAAGG TGGAGCAAA ACACCTCTG 1920
GTCTATAAG GACACCCCAA ATTCCAGAA GTCTGATCAT CTGGCATCC CTCTGGCCT 1980
TGGCTTGTAT TCTTGCAGTT TGCAATGCG TCAACAGTGC AAGAAGGTGT GGGCAGAAGA 2040
AAAGCTAGT GATCAACAGT GGCAATGGAG CTGTGGAGGA CAGAAAGCCA AGTGAATCA 2100
ACGGAGAGGC CAGCAAGTCT CAGGAAATGG TGCATTGGT GAACAAGGAG TCGTCAGAAA 2160
CTCCAGACCA GTTTATGACA GCTGATGAGA CAAGGAACCT GCAGAAATGT GACATGAAGA 2220
TTGGGGTGTG ACACCTACAC CATTATCTTG GAAAGAAACA ACCGTTGGAA ACATAACCAT 2280
TACAGGGAGC TGGGACACTT AACAGATGCA ATGTGCTACT GATTGTTTCA TTGCGAATCT 2340
TTTTTAGCAT AAAATTTTCT ACTCTTAAAA AAAAAA AAAA

```

Seq ID NO: 132 Protein sequence:
Protein Accession #: AAH04372

25
30
35
40

```

1 11 21 31 41 51
MDKFWHAAM GLCLVPLSLA QIDLNITCRF AGVFRVKEKG RYSISRTEAA DLCKAFNSTL 60
PTMAQMEKAL SIGFETCRYG PIEGHVVIPI IHPNSICAA NTGVVILTSN TSQYDTCYFN 120
ASAPPEEDCT SVTDLNPAFD GPITITIVNR DSTRYVQKGE YRTNPEDIYP SNPTDDVSS 180
GSSSERSSST GGYIFYTST VHPIDEDSP WITDSTRIP ATSTSSNTIS AGWEPNEENE 240
DERDRHLSFS GSGIDDDDEF ISSTISTTFR AFDHTKQND WTQWNPFSN PEVLQTTR 300
MTDVRNGTTF AYGNNWPEA HPPLIHHEHH EEBETPHSTS TIQATPSSTT EETATQKEQW 360
FGNRWHEGYR TGPREDSSH TGTAAASHT SHPMQGRTP SPEDSSWTF FNPISHPMGR 420
GHQAGRRMDM DSSHSTTLQP TANPNTGLVE DLDRTGPLSM TTQSSNSQSF STSHGLEED 480
KDHPTTSLT SSNRNDVTGG RRDPNHSEGS TTLLEGYTS YPHTKESRTF IPVTSAKTGS 540
FGVTAVTVGD SNSNVNRSLS GDQDTFHPSG GSHTTGESSES DGHSHGSQEG GANTTSGPIR 600
TPQIPEWLI LALLALALI LAVCIAVNSR RRCQKKKLV INSGNGAVED RKPSGLNGEA 660
SKSQEMVHLV NKESSETPDQ FMTADETRNL QNVDMKIGV

```

Seq ID NO: 133 DNA sequence
Nucleic Acid Accession #: NM_002882
Coding sequence: 150-755

45
50
55
60

```

1 11 21 31 41 51
CGAGGTTCCG GTCTGGGGC GGAGGGAAGA GCGGGCGGGC GGGAGGCGCC GGCGCCAGAC 60
CGGAGGGAAG GGAGCTACGA GTAGCCGCGC AGAGGCGCGC GAGCCAGCGA CGACCGACCC 120
AGCCGAGCCG CCGCGCGCGC CGCGCCCGCA TGGCGGCGCG CAGGACACT CATGAGGACC 180
ATGATACTTC CACTGAGAAT ACAGAGAGT CCAACCATGA CCCTCAGTTT GAGCCATAG 240
TTCTCTTCC TGAGCAAGAA ATTAACAC TGAAGAAGA TGAAGAGGAA CTTTTTAAAA 300
TGGGGGCAAA ACTGTTCGA TTTGCCTCTG AGAACGATCT CCCAGATGG AAGGAGCGAG 360
GCACTGTGTA CGTCAAGCTC CTGAAGCACA AGGAGAAAGG GGCCATCCGC CTCTCATGC 420
GGAGGACAA GACCTGAAG ATCTGTGCA ACCACTACAT CAGCCGATG ATGGAGCTGA 480
AGCCCAACGC AGGTAGCGAC CGTGCTCGG TCTGGAACAC CCAAGCTGAC TTCGCCGAC 540
AGTGCCCAAA GCCAGAGCTG CTGGCCATCC GCTTCTGAA TGCTGAGAA GCACAGAAAT 600
TCAAAACAAA GTTTGAAGAA TGCAGGAAG AGATCGAAGA GAGAGAAAG AAAGCAGGAT 660
CAGGCAAAA TGCATATGC GAAGAAAGTGG CGGAAAGCT AGAAGCTCTC TCGGTGAAG 720
AGGAGACCAA GGAGATGCT GAGGAGAAGC AATAATCGT CTATTTTAT TTTCTTTCC 780
TCTCTTCTT TTCTTTTTT TAAAAAATT TACCCTGCC CTCTTTTTCG GTTTGTTTTT 840
ATCTTTTCT TTTTCAAGG GACGTTATAT AAAGAACTGA ACTC

```

Seq ID NO: 134 Protein sequence:
Protein Accession #: NP_002873

65
70

```

1 11 21 31 41 51
MAAAKDTHEH HDTSTENTDE SNHDPQEPFI VSLPEQEIKT LEEDEEELFK MRKLFRFAS 60
ENDLPWEKER GTGDVKLLKH KEKGAIKLLM RDKTLKICA NHYITPMEL KPNAGSDRAW 120
VWNTHADPAD ECPKPELLAI RFLNAENAQK FTKPFEECRK EIEEREKAG SGKNDHAEKV 180
AEKLEALS VK EETKEDAEK Q

```

Seq ID NO: 135 DNA sequence
Nucleic Acid Accession #: NM_000077.2
Coding sequence: 277-742

75
80
85

```

1 11 21 31 41 51
CCCAACCTGG GCGGACTTCA GTGTGCCAC ATTGCTAAG TGCTCGAGT TAATAGCACC 60
TCTCCGAGC ACTGCTCAC GGGCTCCCT TGCTGGAAG GATACCGCG TCCCTCCAGA 120
GGATTGAGG GACAGGGTCG GAGGGGGCTC TTCCGCCAGC ACCGAGGAA GAAAGAGGAG 180
GGCTGCGCTG GTTACAGAGG GTTGGGGCGG ACCGCGTGG CTGCGGGCTC GCGGAGAGG 240
GGAGACAAA CAGCGGGCGG CGGGAGCAG CATGGAGCG GCGGCGGGA GCAGCATGGA 300
GCCTTGGCT GACTGGCTGG CCACGGCCGC GCGCGGGGT CGGTGAGAG AGGTGCGGGC 360
GCTGCTGAG GCGGGGGCGC TGCCCAACGC ACCGAATAGT TACGATCGA GGCCATCCA 420
GGTCATGAT ATGGGACCG CCCGAGTGGC GGAGCTGCTG CTGCTCCAG GCGCGGAGCC 480

```

CAACTGCGCC GACCCGCGCA CTCTCAACCC ACCCGTGCAC GACGCTGCCC GGGAGGGCTT 540
CTTGGACAGC CTGGTGGTGC TGCACCGGGC CGGGGCGCGG CTGGACGTGC GCGATGCTTG 600
GGGCGCTCTG CCGCTGGACC TGGCTGAGGA GCTGGGCCAT CGCGATGTGC CACCGTACCT 660
GCGCGCGGCT GCGGGGGGCA CCAGAGGCAG TAACCATGCC CGCATAGATG CCGGGGAAGG 720
TCCCTCAGAC ATCCCGGATT GAAAGAACCA GAGAGGCTCT GAGAAACCTC GGGAAACTTA 780
GATCATCAGT CACCGAAGGT CCTACAGGCG CACAACCTGCC CCGGCCACAA CCCACCCGCG 840
TTTGTAGATT TTCAATTAGA AAATAGAGCT TTTAAAAATG TCCTGCTTTT TAACGTAGAT 900
ATATGCTCTC CCCCACTACC GTAAATGTCC ATTTATATCA TTTTATATAT ATTCTTATAA 960
AAATGTAAAA AAGAAAAACA CCGCTTCTGC CTTTTCACCTG TGTGGAGATT TTCTGGAGTG 1020
AGCACTCAGC CCTAAGCGC ACATTCATGT GGGCAATTCT TCGAGAGCTC GCAGCCTCCG 1080
GAAGCTGTGC ACTTCATGAC AAGCATTTTG TGAAGTAGGG AAGCTCAGGG GGGTTACTGG 1140
CTTCTCTTGA GTCACACTGC TAGCAATGCG CAGAACCAAA GCTCAATAAA AAATAAAATA 1200
ATTTTCATTC ATTCACCTC

Seq ID NO: 136 Protein sequence:
Protein Accession #: NP_000068.1

1 11 21 31 41 51
MEPAAGSSME PSADWLATAA ARGRVEEVRA LLEAGALPNA PNSYGRRPDQ VMMSGARVA 60
ELLLLHGAEP NCADPATLTR PVHDAAREGF LDTLVVLHRA GARLDVRDAW GRLPVDLAE 120
LGHDRVARYL RAAAGGTRGS NHARIDAABG PSDIPD

Seq ID NO: 137 DNA sequence
Nucleic Acid Accession #: NM_058196.1
Coding sequence: 104-421

1 11 21 31 41 51
TGTGTGGGGG TCTGCTTGGC GGTGAGGGGG CTCTACACAA GCTTCCTTTC CGTCATGCGG 60
GCCCCACCCC TGCTCTGAC CATCTGTGTC TCTCTGGCAG GTCATGATGA TGGGCAGCGC 120
CCGAGTGGCG GAGCTGCTGC TGCTCCACGG CGCGGAGCCC AACTGCGCCG ACCCGGCCAC 180
TCTACCCGGA CCGGTGCAGC AGCTGCCCGG GGAGGGCTTC CTGGACACGC TGGTGGTGCT 240
GCACCGGGCC GGGGCGCGGC TGGACGTGCG CGATGCCTGG GGGCGTCTGC CCGTGGACCT 300
GGCTGAGGAG CTGGGCCATC GCGATGTGCG ACGGTACCTG CGCGCGGCTG CCGGGGGCAC 360
CAGAGGCAGT AACCATGCCC GCATAGATGC CGCGGAAGGT CCGTCAGACA TCCCGGATTG 420
AAGAACCAG AGAGGCTCTG AGAAACCTCG GGAACCTTAG ATCATCAGTC ACCGAAGGTC 480
CTACAGGGCC ACAACTGCC COGCCACAA CCACCCCGCT TCGTAGTTT TCATTAGAA 540
AATAGAGCTT TTAATAATGT CCTGCCCTTT AACGTAGATA TAAGCCTTCC CCCACTACCG 600
TAAATGTCCA TTTATATCAT TTTTATATA TTCTTATAAA AATGTAAAAA AGAAAAACAC 660
CGCTTCTGCC TTTTCACTGT GTTGGAGTTT TCTGGAGTGA GCATCAACGC CTTAAGCGCA 720
CATTCATGTG GGCATTTCCT GCGAGCCTCG CAGCCTCCGG AAGCTGTGCA CTTCTATGCA 780
AGCATTTTGT GAACTAGGGA AGCTCAGGGG GGTACTGCG TCTCTTGA TCACTGCT 840
AGCAATGGC AGAACCAAG CTCAATAAA AATAAATAA TTTTCATTCA TCACTC

Seq ID NO: 138 Protein sequence:
Protein Accession #: NP_478103.1

1 11 21 31 41 51
MMMSGARVAE LLLLHGAEPN CADPATLTRP VHDAAREGFL DTLVVLHRAQ ARLDVRDAWG 60
RLPVDLAEEL GHDRVARYLR AAGGTRGSN HARIDAABG SDIPD

Seq ID NO: 139 DNA sequence
Nucleic Acid Accession #: NM_058197.1
Coding sequence: 272-684

1 11 21 31 41 51
CCCAACCTGG GCGGACTTCA GGTGTGCCAC ATTGCTAAG TGCTGGAGT TAATAGCACC 60
TCCTCCGAGC ACTGCTCAC GGCCTCCCTT TGCTGGAAA GATACCGCGG TCCCTCCAGA 120
GGATTGTAGG GACAGGGTCG GAGGGGGCTC TTCCGCCAGC ACCGGAGGAA GAAAGAGGAG 180
GGGCTGGCTG GTCACCAAGG GGTGGGGCGG ACCGCTGCG CTGCGCGGCT GCGGAGAGGG 240
GGAGAGCAGG CAGCGGGCGG CCGGGAGCAG CATGGAGCGG CGCGCGGGA GCAGCATGGA 300
GCCGCGGGCG GGGAGCAGCA TGGAGCCTTC GGCTGACTGG CTGGCCACGG CCGCGGCCCG 360
GGGTGCGGTA GAGGAGGTGC GGGCGTCTCT GGAGGCGGGG GCGCTGCCCA ACGCACCGAA 420
TAGTTACGCT CCGAGGCGCA TCCAGGTGGG TAGAAGGTCT GCAGCGGGAG CAGGGGATGT 480
CGGGCGACTC TGGAGGACGA AGTTTGACGG GGAATTGGAA TCAGTAGCG CTTCGATTCT 540
COGAAAAAGG GGGAGGCTTC CTGGGGAGTT TTCAGAAGGG GTTTGTAATC ACAGACCTCC 600
TCCTGGCGAC GCCCTGGGGG CTTGGGAAAC CAAGGAAGAG GAATGAGGAG CCACGCGCGT 660
ACAGATCTCT CGAATGCTGA GAAGATCTGA AGGGGGGAAC ATATTGTAT TAGATGGAAG 720
TCATGATGAT GGGCAGCGCC CGAGTGGCGG AGCTGCTGCT GCTCCACGGC GCGGAGCCCA 780
ACTGCGCGCA CCGCGCCACT CTCACCGAC CCGTGCACGA CGCTGCCCGG GAGGGCTTCC 840
TGGACACGCT GGTGGTCTG CACCGGGCGG GGGCGCGGCT GGAGCTGCGC GATGCTGGG 900
GCGTCTGCCC CGTGGACCTG GCTGAGGAGC TGGGCCATCG CGATGTGCGA CCGTACCTGC 960
GCGCGGCTGC GGGGGGCACC AGAGGCASTA ACCATGCCCG CATAGATGCC GCGGAAGGTC 1020
CCTCAGACAT CCGGATTGA AAGAACCAGA GAGGCTCTGA GAAACCTCGG GAACTTAGAT 1080
CATCAGTAC CGAAGGTCT ACAGGGCCAC AACTGCCCCC GCCACAACCC ACCCGCTTTT 1140
CGTAGTTTTC ATTTAGAAAA TAGAGCTTTT AAAAATGTCC TGCTTTTAA CGTAGATATA 1200
TGCTTCCCC CACTACCGTA AATGTCCATT TATATCATTT TTTATATATT CTTATAAAAA 1260
TGTAATAAAG AAAACACCG CTTCTGCCCT TTCACTGTGT TGGAGTTTTC TGGAGTGAGC 1320
ACTCACGCC TAAGCGCAC TTCATGTGGG CATTCTTTCG GAGCTCGCA GCCTCCGGA 1380
GCTGTGACT TCATGACAAG CATTTTGTGA ACTAGGGAAG CTCAGGGGGG TACTGCTT 1440
CTCTTGAGTC ACCTGCTAG CAAATGCCAG AACCAGAGCT CAATAAAAA TAAATAATT 1500

TTCATTCAATT CACTC

Seq ID NO: 140 Protein sequence:
Protein Accession #: NP_478104.1

1	11	21	31	41	51	
MEPAAGSSME	PAAGSSMEPS	ADWLATAAAR	GRVEEVRLAL	EAGALPNAPN	SYGRRPIQVG	60
RRSAAGAGDG	GRLWRTKFG	ELESGSASIL	RKKGRLPGEF	SEGVGNHRPP	PGDALGAWET	120
KEEE						

Seq ID NO: 141 DNA sequence
Nucleic Acid Accession #: NM_058195.1
Coding sequence: 163-684

1	11	21	31	41	51	
CCTCCCTACG	GGCGCCTCCG	GCAGCCCTTC	CCGCGTGCGC	AGGGCTCAGA	GCCGTTCGGA	60
GATCTTGGAG	GTCCGGGTGG	GAGTGGGGGT	GGGGTGGGGG	TGGGGGTGAA	GTTGGGGGGC	120
GGGCGGCTC	AGGGAAGCG	GGTGGCGGCC	TGCGGGGGGG	AGATGGGCAG	GGGGCGGTGC	180
GTGGGTCCCA	GTCTGCAGTT	AAGGGGGCAG	GAGTGGCGCT	GCTCACTCT	GGTGCCAAAG	240
GGCGGGCGAG	CGGTGCGCGA	GCTCGGCCCT	GGAGGCGCGG	AGAACATGGT	GCGCAGGTTC	300
TTGGTGACCC	TCCGATTTCG	GCAGCGGTGC	GGCCCGCCGC	GAGTGAGGGT	TTTCTGTGTT	360
CACATCCCGC	GGCTCACGGG	GGAGTGGGCA	GCGCCAGGGG	CGCCCGCCGC	TGTGGCCCTC	420
GTGCTGATGC	TACTGAGGAG	CCAGCGTCTA	GGGCAGCAGC	CGCTTCTTAG	AAGACCAGGT	480
CATGATGATG	GGCAGCGCCC	GAGTGGCGGA	GCTGCTGCTG	CTCCACGGCG	CGGAGCCCAA	540
CTGCGCGGAC	CCGCGCACTC	TCACCGGACC	CGTGCAAGAC	GCTGCCGGGG	AGGGCTTCTT	600
GGACAGCTCG	GTGTGCTGCG	ACCGGGCCGG	GGCGCGGCTG	GACGTGCGCG	ATGCTTGGGG	660
CCGTCTGCCC	GTGGACCTGG	CTGAGGAGCT	GGGCCATCGC	GATGTGCGAC	GGTACCTGGG	720
CGCGGCTGGG	GGGGGCACCA	GAGGCAGTAA	CCATGCCCGC	ATAGATGCGG	CGGAAGGTCC	780
CTCAGACATC	CCCCATTGAA	AGAACCAGAG	AGGCTCTGAG	AAACCTCGGG	AAACTTAGAT	840
CATCAGTCA	CGAAGTCTCT	ACAGGGCCAC	AACTGCCCCC	GCCACAAACC	ACCCCGCTTT	900
CGTAGTTTTT	ATTAGAAAAA	TAGAGCTTTT	AAAAATGTCC	TGCTTTTAA	CGTAGATATA	960
TGCTTCCCCC	CAGTACCGTA	AATGTCCATT	TATATCATTT	TTTATATATT	CTTATAAAAA	1020
TGTAATAAAG	AAAAACACCG	CTTCTGCCCT	TTCAGTGTGT	TGGAGTTTTT	TGGAGTGAGC	1080
ACTCAGCCCC	TAAGCGCACA	TTCATGTGGG	CATTCTTTGC	GAGCCTCGCA	GCCTCCGGAA	1140
GCTGTGAGCT	TCATGACAAG	CATTTTGTGA	ACTAGGGAAG	CTCAGGGGGG	TTACTGGCTT	1200
CTCTTGATGC	ACACTGCTAG	CAATGGCAG	AACCAAGACT	CAATATAAAA	TAAATAAATT	1260
TTCAATCAATT	CACTC					

Seq ID NO: 142 Protein sequence:
Protein Accession #: NP_478102.1

1	11	21	31	41	51	
MGRGRCVGPS	LQLRGQEWRC	SPLVPKGGAA	AAELGPGGGE	NMVRRLVTL	RIRRACGPPR	60
VRVVFVHPR	LTGEWAAPGA	PAVALVLM	LRSRLGQQP	LPRRPDHDDG	QRPSGGAAAA	120
PRRGAQLRRP	RHSHPRARR	CPGGLPGHAG	GAAPGRGAAG	RARCLGPSAR	GPG	

Seq ID NO: 143 DNA sequence
Nucleic Acid Accession #: NM_018131
Coding sequence: 412..1107

1	11	21	31	41	51	
GAAATTGCAC	ACTTAAGAC	ATCAGTGGAT	GAAATCACAA	GTGGGAAAGG	AAAGCTGACT	60
GATAAAGAGA	GACAGAGACT	TTTGGAGAAA	ATTGAGTCC	TTGAGGCTGA	GAAGGAGAAG	120
AATGCTTATC	AACTCACAGA	GAAGACAAA	GAAATACAGC	GACTGAGAGA	CCAATGAAAG	180
GCCAGATA	GTACTACCG	ATTGCTTGA	CAGCTGGAAG	AGACACAGAG	AGAAGGAGAA	240
AGGAGGGAGC	AGGTGTTGAA	AGCCTTATCT	GAAGAGAAAG	ACGTATTGAA	ACAACAGTTG	300
TCTGCTGCAA	CCTCAGCAAT	TGCTGAACTT	GAAAGCAAAA	CCAATACACT	CCGTTTATCA	360
CAGACTGTGG	CTCCAACCTG	CTTCAACTCA	TCAATAAATA	ATATTATGTA	AATGGAAATA	420
CAGCTGAAAG	ATGCTCTGGA	GAAAAATCAG	CAGTGGCTCG	TGTATGATCA	GCAGCGGGAA	480
GTCTATGTAA	AAGGACTTTT	AGCAAGATC	TTTGAGTTGG	AAAAGAAAAA	GGAAACAGCT	540
GCTCATTAC	TCCCAAGCA	GACAAAAAAG	CCTGAATCAG	AAGGTTATCT	TCAAGAAGAG	600
ARGCAGAAAT	GTTACAACGA	TCTCTTGGCA	AGTGCAAAAA	AAGATCTTGA	GGTTGAACGA	660
CAAAACATAA	CTCAGCTGAG	TTTTGAACCTG	AGTGAATTTC	GAAGAAAAATA	TGAAGAAACC	720
CAAAAGAAAG	TTCAAAATTT	AAATCAGCTG	TTGTATTAC	AAAGAAAGGC	AGATGTGCAA	780
CATCTGGAAG	ATGATAGGCA	TAAACACAG	AAGATACAAA	AACTCAGGGA	AGAGAATGAT	840
ATTGCTAGGG	AAAAACTTGA	AGAAGAGAAG	AAGAGATCCG	AAGAGCTCTT	ATCTCAGGTC	900
CAGTCTCTTT	ACACATCTCT	GCTAAAGCAG	CAAGAGAAAC	AAACAAGGGT	AGCTCTGTGG	960
GAACAACAGA	TGCAGGCATG	TACTTTAGAC	TTTGAAAAATG	AAAAACTCGA	CCGTCAACAT	1020
GTGCAGCATC	AATTGCATGT	AATTCTTAAG	GAGCTCCGAA	AGCAAGAA	AAATAACACA	1080
GTGGAATCT	TTGAAACAGC	TTCAATGAGTT	TGCCATCACA	GAGCCATTAG	TCACITTTCCA	1140
AGGAGAGACT	GAAGACAGAG	AAAAAGTTGC	CGCTCACCA	AAAAGTCCCA	CTGCTGCACT	1200
CAATGGAAGC	CTGGTGGAAAT	GTCCCAAGTG	CAATATACAG	TATCCAGCCA	CTGAGCATCG	1260
CGATCTGCTT	GTCCATGTGG	AATACTGTTC	AAAGTAGCAA	AATAAGTATT	TGTTTGTATA	1320
TTAAAGATT	CAATACTGTA	TTTCTGTGTA	GCTGTGGGGC	ATTTTGAATT	ATATATTCTA	1380
CATTTTGAT	AAAACGCTT	ATCTACCTTT	GACACTCCAG	CATGCTAGTG	AATCATGTAT	1440
CTTTTAGGCT	GCTGTGCAAT	TCTCTGGCA	GTGATACCTC	CCTGACATGG	TTTATCATCA	1500
GGCTGCAATG	ACAGAAATGTG	GTGAGCAGCG	TCTACTGAGA	TACTAACATT	TTGCACTGTC	1560
AAAACTACTG	GTGAGGAAAA	GATAGCTCAG	GTTATTGCTA	ATGGGTTAAT	GCACCAGCAA	1620
GCATAATATT	TTATGTTTCG	GGGGTTTGA	AAAATCAAAG	ATAATTAACC	AAGGATCTTA	1680
ACTGTGTTCC	CATTTTITAT	CCAAGCACTT	AGAAAACTTA	CAATCTAAT	TTTGTGTCC	1740
ATTGTAAAG	GGTGTGATA	GATACTATT	TTTTTTCATA	TTGTATAGCG	GTTATTAGAA	1800

WO 02/086443

PCT/US02/12476

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

```
AAGTTGGGGA TTTTCTTGAT CTTTATTGCT GCTTACCATT GAAACTTAAC CCAGCTGTGT 1860
TCCCAACTC TGTCTGCGC ACGAAACAGT ATCTGTTTGA GGCATAATCT TAAGTGGCCA 1920
CACACAATG TTTCTCTTAT GTTATCTGGC AGTAACTGTA ACTTGAATTA CATTAGCACA 1980
TTCTGCTTAG CTAAATTTGT TAAAATAAAC TTAAATAAAC CCATGTAGCC CTCTCATTTG 2040
ATTGACAGTA TTTTAGTTAT TTTTGGCATT CTTAAAGCTG GGCAATGTAA TGATCAGATC 2100
TTTGTGTTG TCAGACAGTA TTTTATACA TGCTTTTGTG AAACCAAAAA CTTTAAATTT 2160
TCTTCAGGTT TCTTAACATG CTTACCACGT GGCTACTGTA AATGAGAAAA GAATAAAATT 2220
ATTTAATGTT TT
```

Seq ID NO: 144 Protein sequence:
Protein Accession #: NP_060601

1 11 21 31 41 51
MEIQLKDALE KNQQLVYDQ QREVVYKGLL AKIFELEKKT ETAHSLPQQ TKKPESEGYL 60
QEEKQKCYND LLASAKKDLR VERQTITQLS FELSEPRRKY ESTQKEVHNL NOLLYSQRRR 120
DVQHELEDRH KTEIKQLRE ENDIARGKLE EEKRRSEELL SQVQLYLSL LKQEEQTRV 180
ALLEQQMQAC TLDPENKELD RQHVQHLHV ILKELRKARK NNTVGILETA S

Seq ID NO: 145 DNA sequence
Nucleic Acid Accession #: NM_001168
Coding sequence: 50..478

1 11 21 31 41 51
CCGCCAGATT TGAATCGCGG GACCCGTGGC CAGAGGTGGC GGCGCGGCA TGGGTGCCCC 60
GACGTTGCCCC CCTGCTGGCC AGCCCTTTCT CAAGGACCAC CGCATCTCTA CATTCAAGAA 120
CTGGCCCTTC TTTGAGGGCT GCGCCTGCAC CCGGAGCGG ATGGCCGAGG CTGGCTTCAT 180
CCACTGCCCC ACTGAGAACG AGCCAGACTT GGCCCACTGT TTCTTCTGCT TCAAGGAGCT 240
GGAAGGCTGG GAGCCAGATG ACGACCCCAT AGAGGAACAT AAAAAGCATT CGTCCGGTTG 300
CGCTTTCCTT TCTGTCAAGA AGCAGTTTGA AGAATTAACC CTTGGTGAAT TTTTGAAACT 360
GGACAGAGAA AGAGCCAAGA ACAAATTTGC AAAGGAAACC AACATAAGA AGAAGAATT 420
TGAGGAAACT GCGAAGAAAG TGCCCGGTGC CATCGAGCAG CTGGCTGCCA TGGATTGAGG 480
CCTCTGGCCG GAGCTGCTCT GTCCAGAGT GGCTGCACCA CTTCCAGGGT TTATTCCTTG 540
GTGCCACCAG CCTTCTGTG GGCCTCTTAG CAATGTCTTA GGAAAGGAGA TCAACATTTT 600
CAAATTAGAT GTTCAACTG TGCTCCGTGT TTGTCTTGAA AGTGGCACA GAGTGCTTC 660
TGCCTGTGCA GCGGGTGCTG CTGGTAACAG TGGCTGCTTC TCTCTCTCTC TCTCTTTTTT 720
GGGGGCTCAT TTTTGTGTTT TTGATTCCTG GGCTTACCAG GTGAGAAGTG AGGGAGGAAG 780
AAGGCACTGT CCTTTTGTCT AGAGCTGACA GCTTTGTTGG CGTGGGCAGA GCCTTCCACA 840
GTGAATGTGT CTGGACCTCA TGTGTGTGAG GCTGTACAG TCCTGAGTGT GGACTTGGCA 900
GGTCCCTGTT GAATCTGAGC TGCAGGTTCC TTATCTGTCA CACCTGTGCC TCCTCAGAGG 960
ACAGTTTTTT TGTGTGTGTT TTTTGTGTT TTTTGTGTT GGTAGATGCA TGACTTGTGT 1020
GTGATGAGAG AATGGAGACA GAGTCCCTGG CTCTCTACT GTTAAACAAC ATGGCTTTCT 1080
TATTTGTGTT GAATTTGTTA TTCAAGAAAT AGCACAAACT ACAATTAAAA CTAAGCACA 1140
AGCCATTCTA AGTCATTGGG GAAACGGGGT GAACTTCAGG TGGATGAGGA GACAGAAATAG 1200
AGTATAGAGA AGGCTCTGGC AGATACTCCT TTTGCCACT CTGTGTGATT AGACAGGCC 1260
AGTGAGCCCG GGGGCACATG CTGGCCGCTC CTCCCTCAGA AAAAGGCAGT GGCTAAATC 1320
CTTTTAAAT GACTTGGCTC GATGCTGTGG GGGACTGGCT GGGCTGCTG AGGCCGTGTG 1380
TCTGTGAGCC CAACCTTCAC ATCTGTACAG TTCTCCACAC GGGGAGAGA CGCAGTCCGC 1440
CCAGTCCCC GCTTCTTTG GAGGCAGCAG CTCGCCGAGG GCTGAAGTCT GCGCTAAGAT 1500
GATGGATTGT ATTCTGCCCTC CTCCCTGTCA TAGAGCTGCA GGGTGGATTG TTACAGCTTC 1560
GCTGGAACC TCTGAGGTC ATCTCGGCTG TTCTGAGAA ATAAAAGCC TGTCATTTT

Seq ID NO: 146 Protein sequence:
Protein Accession #: NP_001159

1 11 21 31 41 51
MGAPTLPPAW QPFLKDHRS TFKNWPFLG CACTPERMAE AGFIHCPTEN EPDLAQCFPC 60
FKELEGWEPD DDPIEHHKHH SSGCAFLSVK XQFEELTLGE FLKLDREKAR NKIAKETNNK 120
KKEFEETAKK VRRRAISQLAA MD

Seq ID NO: 147 DNA sequence
Nucleic Acid Accession #: NM_014176.1
Coding sequence: 127-720

1 11 21 31 41 51
GCGCGCAGCG CTGGTACCCC GTTGGTCCGC GCGTTGCTGC GTTGTGAGGG GTGTCAGCTC 60
AGTGCAATCCC AGGCAGCTCT TAGTGTGGAG CAGTGAACCTG TGTGTGGTTC CTCTACTTGT 120
GGGATCATCG AGAGAGCTTC ACGTCTGAAG AGAGAGCTGC ACATGTTAGC CACAGAGCCA 180
CCCCAGGCA TCACATGTTG GCAAGATAAA GACCAATAGG ATGACCTGCG AGCTCAAAAT 240
TTAGGTGAGG CCAACACACC TTATGAGAAA GGTGTTTTTA AGCTAGAAGT TATCATTTCT 300
GAGAGGTACC CATTGAACCT TCCTCAGATC CGATTCTCTA CTCAATTATA TCATCAAAAT 360
ATTGATTCTG CTGGAAGGAT TTGTCTGGAT GTTCTCAAT TGCCACCAAA AGTGTCTTGG 420
AGACCATCCC TCAACATCGC AACTGTGTTG ACCTCTAATC AGCTGCTCAT GTCAAGACC 480
AACCTTGATG ACCCGTCTAT GGTGACATA TCCTCAGAAT TTAATATAA TAAGCCAGCC 540
TTCTCAAGA ATCCAGACA GTGACAGAG AAGCATGCAA GACAGAAACA AAAGGCTGAT 600
GAGGAAGAGA TGCTGTGATA TCTACAGAG GCTGGTGAAT CCAGAGTACA CAACTCAACA 660
CAGAAAAGGA AGGCCAGTCA GCTAGTAGGC ATAGAAAAGA AATTTCATCC TGATGTTTAG 720
GGGACTTGTG CTGGTTCATC TTAGTTAATG TGTCTTTTGC CAAGGTGATC TAAGTTCCTT 780
ACCTGAATT TTTTATAAA TATATTGAT GACATAATTT TTGTGATGTT TATTTATCTT 840
GTACATATGT ATTTTGAAT CTTTAAACC TGAAAAATA ATAGTCATTT AATGTTGAAA 900

AAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 148 Protein sequence:
Protein Accession #: NP_054895.1

1	11	21	31	41	51	
MQRASRLKRE	LHMLATEPPP	GITCWQDKDQ	MDDLRAQILG	GANTPYEKGV	FKLEVIIPER	60
YPFEPQIRP	LTPIYHPNID	SAGRICLDVL	KLPPKGAWRP	SLNIATVLT	IQLLMSEPNP	120
DDPLMADISS	EFKYNKPAFL	KNARQWTEKH	ARQQRKADSE	EMLDNLPEAG	DSRVHNSQK	180
RKASQLVGIE	KKFHPDV					

Seq ID NO: 149 DNA sequence
Nucleic Acid Accession #: NM_003812
Coding sequence: 224-272

1	11	21	31	41	51	
TCCTCTGCT	CCGCCCCCGG	GAGTGGCTGC	GAGGCTAGGC	GAGCCGGGAA	AGGGGGCGCC	60
GCCCCAGCCC	GAGCCCGCGG	CCCCGTGCC	CGAGCCCGGA	GCCCCCTGCC	CGCGGCGGCA	120
CCATGCGCGG	CGAGCCCGGG	TGACCGGCTC	CGCCCGCGGC	CGCCCCGCGC	CTAGCCCGGC	180
GCTCTGCGCG	GCCACACGGA	GCGGCGCCCG	GGAGCTATGA	GCCATGAAGC	CGCCCGGCGC	240
CAGCTCGCGG	CAGCCGCCCG	TGGCGGGCTG	CAGCCTTGCC	GGCGCTTCCT	GCGGCCCCCA	300
ACGCGGCCCC	GCGGCGCTCG	TGCTCGCCAG	CGCCCGCGGC	CGCACGCGGC	CCTGCGCGCT	360
GCTTCTCGTC	CTTCTCTGCT	TGCTCCGCT	CGCCGCGCTG	TCCCGGCCCC	GCGCTTGGGG	420
GGCTGCTGCG	CCGACGCTCT	CGCATTGGAA	TGAAACTGCA	GAAAAAAATT	TGGGAGTCTCT	480
GGCAGATGAA	GACAATACAT	TGCAACAGAA	TAGCAGCAGT	AATATCAGTT	ACAGCAATGC	540
AATGCAGAAA	GAATACACAC	TGCTTCAAG	ACTCATATAT	TACATCAACC	AAGACTCGGA	600
AAGCCCTTAT	CACGTTCTTG	ACACAAAGGC	AAGACACCG	CAAAAACATA	ATAAGGCTGT	660
CCATCTGGCC	CAGCAAGCTT	TCCAGATTGA	AGCCTTCGGC	TCCAAATTCA	TTCTTGACCT	720
CATACAGAAC	AATGGTTTGT	TGTCTTCTGA	TTATGTGGAG	ATTCACTACG	AAAATGGGAA	780
ACCACAGTAC	TCTAAGSGTG	GAGAGCACTG	TTACTACCAT	GGAAGCATCA	GAGGCGTCAA	840
AGACTCCAG	GTGGCTCTGT	CACCTTGCAA	TGGACTTCAT	GGCATGTTTG	AAGATGATAC	900
CTTCTGTAT	ATGATAGAGC	CACCTAGAGC	GGTTCATGAT	GAGAAAAGCA	CAGGTGACCC	960
ACATATAATC	CAGAAAACCT	TGGCAGGACA	GATTTCTAAG	CAATGAAGA	ATCTCACTAT	1020
GGAAAGAGGT	GACCAAGTGC	CTTTCTCTC	TGAATTACAG	TGGTTGAAA	GAAGGAAGAG	1080
AGCAGTGAAT	CCATCAGCTG	GTATATTGGA	AGAAATGAAA	TATTTGGAAC	TTATGATTGT	1140
TAATGATCAC	AAAACGTATA	AGAAGCATCG	CTCTTCTCAT	GCACATACCA	ACAACTTTGC	1200
AAAGTCCGTT	GTCAACCTTG	TGGATTCTAT	TTACAAGGAG	CAGCTCAACA	CCAGGGTTGT	1260
CTGTGTGGCT	GTAGAGAGCT	GGACTGAGAA	GGATCAGATT	GACATCACCA	CCAACCTGTT	1320
GCAGATGCTC	CATGAGTTCT	CAAAATACCG	GCAGCGCATT	AAGCAGCATG	CTGATGCTGT	1380
GCACCTCATC	TGCGGGGTGA	CATTTCACCTA	TAAGAGAAGC	AGTCTGAGTT	ACTTTGGAGG	1440
TGCTGTGTCT	CGCACAAGAG	GAGTTGGTGT	GAATGAGTAT	GGTCTTCCAA	TGGCAGTGGC	1500
ACAAATATTA	TGCGAGAGCC	TGGCTCAAAA	CCTTGGAAATC	CAATGGGAAC	CTTCTAGCAG	1560
AAAGCCAAAA	TGTGACTACA	CAGAATCCTG	GGGTGGCTGC	ATCATGGAGG	AAACAGGGGT	1620
GTCCATCTCT	CGAAAATTTT	CAAAAGTCAG	CATTTTGGAG	TATAGAGACT	TTTTACAGAG	1680
AGGAGGTGGA	GCCTGCCTTT	TCAACAGGCC	AACAAAGCTA	TTTGAGCCCA	CGGAATGTGG	1740
AAATGGATAC	GTGGAAGCTG	GGGAGGAGTG	TGATTGTGGT	TTTCATGTGG	AATGCTATGG	1800
ATTATGTGTT	AGAATATGTT	CCCTCTCCAA	CGGGGCTCAC	TGCAGCGAGC	GGCCCTGCTG	1860
TAACAATACC	TGATGTCTTT	TTACAGCCAG	AGGGTATGAA	TGCGGGATG	CTGTGAACGA	1920
GTGTGATATT	ACTGAATATT	GTACTGGAGA	CTCTGGTCAG	TGCCCAACAA	ATCTTCATAA	1980
GCAGAAGCGA	TATGCATGCA	ATCAAAATCA	GGGCGGCTGC	TACAATGGCG	AGTGCAAGAC	2040
CAGAGACAAC	CAGTGTCACT	ACATCTGGGG	AACAAAGGCT	GCAGGGTCTG	ACAAGTTCTG	2100
CTATGAAAG	CTGAATACAG	AAGGCACTGA	GAAGGGAAC	TGCGGGAAGG	ATGGAGACCG	2160
GTGGATTACG	TGCACCAAA	ATGATGTGTT	CTGTGGATTG	TTACTCTGTA	CCAATCTTAC	2220
TCGAGCTCCA	CGTATTGGTC	AACTTCAGGG	TGAGATCATT	CCAACCTCCT	TCTACCATCA	2280
AGGCGGGTGA	ATTGATGCA	GTGGTGCCCA	TGTAGTTTGA	GATGATGATA	CGGATGTGGG	2340
CTATGTAGAA	GATGGAACGC	CATGTGGGCC	GTCTATGATG	TGTTTAGATC	GGAAGTGCCT	2400
ACAAATTCAA	GGCCTAAATA	TGAGCAGCTG	TCCACTCGAT	TCCAAGGGTA	AAGTCTGTTC	2460
GGGCGATGGG	GTGTGTAGTA	ATGAAGCCAC	CTGCATTTGT	GATTTCACCT	GGGCGAGGAC	2520
AGATTGCACT	ATCCGGGATC	CAGTTAGGAA	CCTTCACCCC	CCCAAGGATG	AAGGACCCAA	2580
GGGTCTCTAT	GCCCAATATC	TCATAATAGG	CTCCATCGCT	GGTGCCATCC	TGGTAGCAGC	2640
TATTTGCTCT	GGGGGCACAG	GCTGGGGATT	TAAAAATGTC	AAGAAGAGAA	GTTTCGATCC	2700
TACTCAGCAA	GGCCCCATCT	GAATCAGCTG	CGCTGGATGG	ACACGCGCTT	GCACTGTITG	2760
ATTCTGGGTA	TGACATACTC	GCAGCAGTGT	TACTGGAATC	ATTAAGTTTG	TAAACAAAAC	2820
CTTTGGGTGG	TAACTACTAC	GGAGCTAAAG	TTGGGGTGAC	AAGGATGGGG	TAAAGAAAAA	2880
CTGTCTCTTT	TGGAAATAAT	GTCAAAGAAC	ACCTTTCACC	ACCTGTCACT	AAACGGGGGA	2940
GGGGGCACAA	GACCATGCTA	TAAAAGAAGC	TGTTCCAGAA	TCTTTTTTTT	TCCCTAATGG	3000
ACGAAGGAAC	AACACACACA	CAAAAATTAA	ATGCAATAAA	GGAATCATTAA	AAAA	

Seq ID NO: 150 Protein sequence:
Protein Accession #: NP_003803

1	11	21	31	41	51	
MKPPGSSSRQ	PPLAGCSLAG	ASCGPQRCPA	GSVPASAPAR	TPPCRLLLV	LLLPLLAASS	60
RPRANGAAAP	SAPHWNETAE	KNLGVLADED	NTLQNNSSSN	ISYSNAMOKE	ITLPSRLIYY	120
INQDSESPYH	VLDYKARHQQ	KHNKAVHLAQ	ASFQIEAFGS	KFILDILINN	GLLSSDYVEI	180
HYENKPYQS	KGGEHCYHNG	SIRGVKDSKV	ALSTCNGLHG	MFEDDTFVYM	IEPLELVHDE	240
KSTGRPHIIQ	KTLAQOYSKQ	MONLTMERGD	QWPLSELQW	LKRRKRAVNP	SRGIFPEEMKY	300
LELMIVNDHK	TYKKHRSSEA	HTNFAKSVV	NLVDSIYKEQ	LNTRVVLVAV	ETWTEKDQID	360
ITINPVQMLH	EFKRYRQRIK	QHADAHVHLIS	RVTFRYKRSS	LSYFGVCVSR	TRGVGVNEYG	420
LPMAVAQVLS	QSLAQNLGIQ	WEPSSRKPKC	DCITESWGGCI	MEETGVSHSR	RFSKCSILEY	480
RDFLQRGGGA	CLFNRPKLP	EPTECGNGYV	EAGEECDCGP	HVECYGLCK	KCSLSNGAHC	540
SDGPCCNNTS	CLFQPRGYEC	RDVNECDIT	EYCTGDSGQC	PPNLEKQDGY	ACNQNGRCY	600
NGBCKTRDNG	QYINGCTKAA	GSDKFCYEKL	NTEGTEKGNC	GKGDGRWIOC	SKHDVFCGFL	660

WO 02/086443

LCTNLTRAPR IGQLQGEIIP TSFYHQGRVI DCSGARVVLDD DTDVGVYVED GTPOGSPMMC 720
LDRKLCIQIA LNMSSCPIDS KGVKVCSEHGV CSNEATCICD FTWAGTDCSI RDPVRNLHPP 780
KDEGPKGPSA TNLIGSIAG AILVAAIVLG GTGHPKKNVK KRRFDPQOG PI

PCT/US02/12476

Seq ID NO: 151 DNA sequence
Nucleic Acid Accession #: NM_023915
Coding sequence: 250-1326

1 11 21 31 41 51
GGCAGGAGG TTTCTTTTC ATGCTTTACC AGAAAAATCCA CTTCCTGCC GACCTTAGTT 60
TCAAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG 120
GTGAATGGAC AGCCAGCCAC CACAATGAAA GAAATCAAAC CAGGAATAAC CTATGCTGAA 180
CCACGCGCTC AATCGTCCCC AAGTGTTCCT TGACACGCAT CTTTGCTTAC AGTGCAATCAC 240
15 AACTGAAGAA TGGGGTTCAA CTTGACGCTT GCAAAATTAC CAAATAACGA GCTGCAOGGC 300
CAAGAGAGTC ACAATTGAGG CAACAGGAGC GACGGGCCAG GAAAGAACAC CACCCCTTAC 360
AATGAATTG ACACAATTGT CTTGCGGTG CTTTATCTCA TTATATTGT GGCAAGCATC 420
TTGCTGAATG GTTTAGCAGT GTGGATCTTC TTCCACATTA GGATAAAAC CAGCTTCATA 480
20 TTCTATCTCA AAAACATAGT GGTTCGAGAC CTCATAATGA CGCTGACATT TCCATTTTCA 540
ATAGTCCATG ATGCAGGATT TGGACCTTGG TACTTCAAGT TTATTCTCTG CAGATACACT 600
TCAGTTTGTG TTATATGCAA CATGTATACT TCCATCGTGT TCCCTGGGCT GATAAGCAAT 660
GATCGCTATC TGAAGGTGGT CAAGCCATTG GGGGACTCTC GGATGTACAG CATAACCTTC 720
ACGAAGGTTT TATCTGTTTG TGTCTGGGTG ATCATGGCTG TTTTGTCTTT GCCAAACATC 780
ATCCTGACAA ATGCTCAGCC AACAGAGGAC AATATCCATG ACTGCTCAA ACTTAAAGT 840
25 CCTTTGGGGG TCAAATGGCA TACGGCAGTC ACCTATGTGA ACAGCTGCTT GTTTGTGGCC 900
GTGCTGTGTA TTCTGATCGG ATGTTACATA GCCATATCCA GGTACATCCA CAAATCCAGC 960
AGGCAATTCA TAAGTCAGT AAGCCGAAAG CGAAACATA ACCAGAGCAT CAGGGTTGTT 1020
GTGCTGTGTT TTTTACCTG CTTTCTACCA TATCACTTGT GCAGAATTCC TTTTACTTTT 1080
AGTCACTTAT ACAGGCTTTT AGATGAATCT GCACAAAAAA TCCTATATTA CTGCAAGAA 1140
30 ATTACATTG CTGTGCTGCG GTGTAATGTT TGCCTGGATC CAATAATTTA CTTTTCATG 1200
TGTAAGTCAT TTCAAGAAG GCTGTTCAA AAATCAAATA TCAGAACCAG GAGTGAAAGC 1260
ATCAGATCAC TGCAAGGTGT GAGAAGATCG GAAGTTGCGA TATATTATGA TTCACTGAT 1320
GTGTAGGCTT TTTATTGTTT GTTGAATCG ATATGTACAA AGTGTAATA AATGTTTCTT 1380
35 TTCATTATCC TTAATAAAAA AA

Seq ID NO: 152 Protein sequence:
Protein Accession #: NP_076404

1 11 21 31 41 51
MGFNLTLAKL PNNELHQGES HNSGNRSDGP GKNTLENEF DTIVLPVLYL IIFVASILLN 60
GLAVNIFPHI RNKTSIFPYL KNIVVADLIM TLTFFPRIVH DAGFGPHYFK PILCRYTSVL 120
FYANMYTSIV PLGLISIDRY LKVVKPPGDS RMYSTITFKV LSVCVWVIMA VLSLPNIILT 180
45 NQPTEDNIH DCSKLKSPFG VKNHTAVTYV NSCLFVAVLV ILIGCYIAIS RYIHKSSRQF 240
ISQSSRRKKH NQSRVUVAV FPTCFPLPYL CRIPFTFSL DRLLDESAQK ILYYCKEITL 300
FLSACNVCLD PIIFYMCRPS PSRRLFKKSIN IRTRESIRS LQSVRRSEVR IYYDYTDV

Seq ID NO: 153 DNA sequence
Nucleic Acid Accession #: D80008.1
Coding sequence: 149-739

1 11 21 31 41 51
GTTGCGGCC AAAGCGCGGA GCGGAGGCGG AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60
CGAAGAGAGT GAGGCGCCGA GAGCCAGAT ACCATTTTGG CGTGAGAGCT GGTGGTTGGC 120
AAGGCCGCGG GAGTGGGAAG CGTCGCGCAT GTTCTGCGAA AAAGCCATGG AACTGATCCG 180
CGAGCTGCAT CGCGCGCCCG AAGGGCAACT GCCTGCCTTC AACGAGGATG GACTCAGACA 240
60 AGTTCTGGAG GAGATGAAGG CTTTGTATGA ACAAACCCAG TCTGATGTGA ATGAAGCAAA 300
GTCAGGTGGA CGAAGTGATT TGATACCAAC TATCAAATTT CGACACTGTT CTCTGTTAAG 360
AAATCGACGC TGCAGCTAGT CATACCTGTA TGACGCGCTG CTTCGGATCA GAGCACTCAG 420
ATGGGAATAT GGTAGCGTCT TGCCAAATGC ATTACGATT CACATGGCTG CTGAAGAAAT 480
GGAGTGGTTT AATAATTATA AAAGATCTCT TGCTACTTAT ATGAGGTAC TGGGAGGAGA 540
TGAAGGTTT GACATTACAC AGGATATGAA ACCACCAAAA AGCCTATATA TTGAAGTCCG 600
65 GTGTCTAATA GACTATGGAG AATTGAAAT TGATGATGGC ACTTCAGTCC TATTAAAAAA 660
AAATAGCCAG CATCTTTTAC CTCGATGGAA ATGTGAGCAG CTGATCAGAC AAGGAGTCCCT 720
GGAGCACATC CTGCTATGAC CATGCGCGGA GGCATTCCA GGCCTTCACT AACTCATGGA 780
CTCCTCTGTA CTCACTCTCT CCACCACTCC CTTCACCTCC CTCCTTGATT TTGAAGCTA 840
TAGACATTGT TTAAGATAAC TAAGAATACT TGGCTAAGAA GTATAATTG TAACTATTA 900
70 AGGACTTTCT TTTTTAATG TTGTACACTA TTCTTCTAC TCTTTTGGG TTTTGGTTTT 960
GTTTGTAGA GACTGTCTCA CTATGTTGCC CAAGCTGGTC TCAAACCTCT GGCCTCAAGC 1020
AGTCTCTCCA CCTTAGCTTC TCAAAGTGTG GAGATCACAG GGTGAGCCA CTGCACCCGG 1080
CCCTACTCC TTTTCTAAT AAGCTGTATC TGTAAATACA GCATTCTAC AGTTGTTACA 1140
GTGTGTTTT TAAATGAAGG TAAACATGGT TACATTTGAA TCTCTTAAAT AAGCAGTCAC 1200
75 TTGGCTGGAC AGGAGAGAGG TAGATCCTGT GTGCTTGTG TTCTGGTCAT GTGTATTGTA 1260
CAAGCTAGAG AGCTGAATTT CTGAGATACA CATTTTCAA TCACATGCAA GTGAAGATGA 1320
TGGTCTGTAG AAATTTTCAG TATATATAAT GTTAAATGAC ATACTAATTT ATCATCTGGC 1380
TATTTGGGAA GGAAGGACAC ACATGGATT TGCAATTC CACCATGGTG GCTGGTGGG 1440
CTGTGGCTA TGGGGTATC ACCAGTATCA CCATTTGGA AGGGGACAGT GAAATTTGGG 1500
80 CAGAGAGAG AACTTTGTAC AGTTTCCCT GAGATTGAGA TTGACTGAAA AGTCACATGA 1560
AGAGTTGATT GTCTTTTAAT GGTATGTTTT AAACAGCTGA CATTTTAAAT TTGATGAAA 1620
TCCAGTTTAT TCGTTTGTTC TTTTATGCTT TGGGTGTTGC ATCCGAGAAA TCTTTTCCCA 1680
TCCCAAGATC ACAATTTTTT TTCCTTTTTA CTCTAGAAG TGTATAAAT TTAAGCTTTA 1740
TACTTTGCTC TATGACCCGT TTTTTTTTTT GTTTTGTGTT GTTTTTCGT TTGTTTCTTT 1800
85 GTTTTGTAGT GAGGCTCTGT TCTGTCAACC AGGCTGGGGT GCAGTGGCGT GATCTTGGCT 1860
CACTGCAATC TCTATCCCTT GGGTTCAGT GATTCTCTTG TCTCAGCTC CCAAGTAGCT 1920
GGGATTACAG GCACAGGCGG CCACGCCCTGG CTAATTTTTG TATTTTATAGT AGAGACAGAG 1980

WO 02/086443

PCT/US02/12476

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

TTTTACCATG TTGGCCAGGC TGGTTTCAAA CTCTGACCT CAAGTGACCC ACCTTGGCCT 2040
 CCCAAAGTTT TGGGATTACA AGTGTGGGCC ACCGGGGCCA GCCTATGATC CATTTTGAAT 2100
 GAATTTTATA TATGGTGCAA GGTGTCAATC CACCTTCACT TTTTCTTGGG AATATAGATA 2160
 TCCAGCTGTT TCACTACCAT TTTTGAAGG GACTGCCCTT TGCTCTATCA CCTTTGCATT 2220
 TTTGTTAAAA AGTAGTTGTC AATGTATATG TGGGTTTATT TCAGSACTCT GTTTTGTTC 2280
 ATTGACCTGT TTTTCTCTCC TGAATGCCAA TACCATATTT GTATGTAGTG TATGTAATTT 2340
 TCTAATAATT CTGAAACAG ATAGTATTAA TGTGTCTAT TTTTGTCTGT GTTTGTATTT 2400
 TTTGTAGAGA TGGGGTTTCA CCGTGTGGC CAGGCTGTGT TGAATCCTG AGCTAAAGCA 2460
 ATACACTTGC CTGCTCTCTC CCATGTGCTG GGATACAGG CGTGAGCCTT GGTGCTGGCC 2520
 CAGTGTACCA CATTTCTTTT TGAGATTGT TTTGGCTATG TTAAGTCTCT TGCTTTTGAT 2580
 GTGAAATTTG GGAACAGCA GGGTGTGGTG GCTTATGCTT GTAATCCTAG AACTTTGGGA 2640
 GGCCTAGATG GGTGGATCAC TTGAGCTCAG GAGTTCAGG CCAGCCCGGG CCTATGGCAA 2700
 AACTCCGTCT CTACAAAAA TAGAAAAAAT TAGCCAGGTG TGGTGGTGCA TGCTGTAGT 2760
 CACAGTTACA CGGCAGGCTG AGGTGGGAGG ATCACTTGA CCCCAGAGGT CAAGACTGCA 2820
 GTGAGCTGAG ATCACCAC TGTACTCCAG CCTGGGTGAC AAAGTGAGAC TCTATCTCAA 2880
 AAAGAAATTA GGATCAATTT GTCATTTTCT ACAACAACA CAACAAAAAC CCTGTGGG 2940
 CACCTTGATT GAGATTGCAT TGAATTTATA TAAACTGTT GGGAGAAATG ACATCTTAAT 3000
 AATATTGAGT CTTCGCGCT ATAAACAAGG TCTGTCTTCC TAGGTATTAA TGTTTGTCT 3060
 TCTATTCTC TTAATAATCT TTTGTAGTT TCAGTGTACA GGTCTACCAT GTGAGCAATT 3120
 CATAGTTTGT ATGCTAAATG GTATTTTAAA ATTTCAAAT CTAACCACT GTTGCTAGTA 3180
 AATAGAAATA CAATTGATG TGAACCTGTA TCCTTCAGCC TTGCTAACT GTGAGTTCTC 3240
 ATGGTGTITT TGTAATATC ATCAACAGTC ATGTGTTCTA TGAATAAGA GTTTTACTCC 3300
 TTC

Seq ID NO: 154 Protein sequence:
 Protein Accession #: BAA11503.1

30
35
40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
 MFCEKAMELI RELHRAPEGO LPAFNEGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIP 60
 TIKFRHCSLL RNRRTVAYL YDRLLRIRAL RWEYGSVLPN ALRFHMAAEE MEWFNMYKRS 120
 LATYMRSLGG DEGLDITQDM KPPKSLYIEV RCLKDYGEFE VDDGTSVLLK KNSQHLPRW 180
 KCBQLIRQGV LEHILS

Seq ID NO: 155 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 149-709

40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
 GTTCGGCGCC AAAGCGCGGA GCGGAGGCCG AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60
 CGAAAGGAGT GAGGCGCCGA GAGCCAGAT ACCATTTTGG: CGTGAGAGCT GGTGGTGGGC 120
 AAGGCGCGCG GAGTGGGAAG CGTCCGCCAT GTTCTGCGAA AAAGCCATGG AACTGATCCG 180
 CGAGCTGCAT CGCGCGCCCG AAGGGCAACT GCCTGCCTTC AACGAGGATG GACTCAGACA 240
 AGTTCGAGAG GAGATGAAAG CTITGTATGA ACAAACCAAG TCTGATGTGA ATGAAGCAAA 300
 GTCAGGTGGA CGAAGTGATT TGATACCAAC TATCAAAATT CGACACTGTT CTCTGTTAAG 360
 AAATCGACGC TGCACGTGAG CATACCTGTA TGACCGCTTG CTTCGGATCA GAGCACTCAG 420
 ATGGGAATAT GGTAGCGTCT TGCCAAATGC ATTACGATT CATATGGCTG CTGAAGAAAT 480
 GGAGTGGTTT AATAATTATA AAAGATCTCT TGCTACTTAT ATGAGGTGAC TGGGAGGAGA 540
 TGAAGGTTTG GACATTACAC AGGATATGAA ACCACCAAAA AGCCTATATA TTGAAGCTGG 600
 ATGCACTGGC GCGATCTCGG CTCACCTGAC AACCTCCACC TCCACAGTTC ACCTCAACTG 660
 CAACCTCCAC CTCCAGGTC CGGTGTCTAA AAGACTATGG AGAATTTGAA GTTGATGATG 720
 GCACTTCAGT CCTATTAAAA AAAAATAGCC AGCACTTTT ACCTCGATGG AAATGTGAGC 780
 AGCTGATCAG ACAAGGATGC CTGGAGCACA TCCGTGTCAT ACCATGCGCC GAGGCACTTC 840
 CAGGCTTCAC TCACTCATG GACTCCTCTG TACTCACTCT CTCACCACT CCCTTCACTC 900
 CCCTCTTTGA TTTTGAAGC TATAGACATT GTTTAAGATA ACTAAGAATA CTGGCTAAG 960
 AAGTATAATT TGCTAACTAT TAAGGACTTT CTTTTTTAA TGTGTACAC TATTCTTCTC 1020
 ACTCTTTTTT GGTTTTGGTT TTGTTTGTGA GAGACTGTCT CACTATGTTG CCCAAGCTGG 1080
 TCTCAAACCT CTGGCTCTCA GCAGTCTCTC CACTTAGCT TCTCAAAGTG TTGAGATCAC 1140
 AGGCGTGAGC CACTGCAACC GGGCCCTACT CCTTTTCTA ATAAGCTGTA TCTGTAATCA 1200
 CAGCATTCCT ACAGTGTGTA CAGTGTGTTT TTTAAATGAA AGTAAACATG GTTACATTG 1260
 AATCTCTTAA ATAAGCAGTC ACTTGGCTGG ACAGGAAGAA GGTAGATCCT GTGTGTCTTG 1320
 TTTTCTGGTC ATGTGTATTG TACAAGCTAG AGAGCTGAAT TTCTGAGATA CACATTTTCA 1380
 AATCACATGC AAGTGAAGAT GATGGTCTGT AGAAATTTTC AGTATATATA ATGTTTAATG 1440
 ACATACTAAT TTATCATCTG GCTATTGGG AAGGAAGGAC ACACATGGAT TTGCACTT 1500
 TCCACCATGG TGGCTGGTGT GGCTTGTGGC TATGGGGTGA TCACCAATAT CACCACTTTG 1560
 GAAGGGGACA GTGAAATTGG GGCTAGAGAA GGAACCTTGT ACAGTTTTC CTGAGATTCA 1620
 GATTGACTGA AAAGTCATAT GAAGAGTTGA TTGTCTTTTA ATGGTATGTT TTAACAGCT 1680
 GACATTTTAA ATTTTATGTA AATCCAGTTT ATTCTGTTGT TCTTTTATGC TTTGGGTGTT 1740
 GCATCCGAGA AATCTTTTCC CATCCCAAGA TCACAAATTT TTTTCTTTT TACTTCTAGA 1800
 AGTGTATAA TTTTAACTT TATACTTGG TCTATGACCC GTTTTTTTT TGTGTTTGT 1860
 TTGTTTTTCT GTTGTGTTCT TTGTTTGGG ATGGAGTCTT GTTCTGTAC CCAGGCTGGG 1920
 GTGCACTGGC GTGATCTTGG CTCACTGCAA TCTCTATCCC CTGGGTTCAA GTGATTTCT 1980
 TGTCTCAGCC TCCCAAGTAG CTGGGATTAC AGGCACAGGC CGCCAGCCTT GGCTAATTTT 2040
 TGTATTTTAA GTAGAGACAG AGTTTATCCA TGTGGCCAG GCTGGTTTCA AACTCCTGAC 2100
 CTCAAGTGAC CCACCTTGGC CTCCCAAAGT TTTGGGATTA CAAGTGTGG CCACCGCGGC 2160
 CAGCCTATGA TCCATTTTGA ATGAATTTT TATATGGTGC AAGGTGTCAA TCCACCTTCA 2220
 CTTTTTCTG GGAATATAGA TATCCAGCTG TTCTACTACC ATTTTGTGAA AGGACTGCCC 2280
 TTTGCTCTAT CACCTTTGCA TTTTGTGTA AAAGTAGTTG TCAATGTATA TGTGGGTTTA 2340
 TTTCAAGACT CTGTTTGTGT CCATTGACCT GTTTTCTCT CCTGAATGCC AATACCATAT 2400
 TTGTATGTAG TGTATGTAAT TTTCTAATAA TTCTTGAAC AGATAGTATT AATGTGTCT 2460
 ATTTTGTCTG TTGTTGTGAT TTTTGTAGA GATGGGGTTT CACGTGTGTC GCGAGGCTGT 2520
 GTTGAACCTC TGAGCTAAAG CAATACACTT GCCTCGTCTT CCCCATGTGC TGGGATTACA 2580
 GCGGTGAGCC TTGGTGTGGG CCGAGTGTAC CACATTTCTT TTTGAGATT GTTTTGGCTA 2640
 TGTTAAGTCC TTTGCTTTTG ATGTGAAATT TGGGAACAGG CAGGGTGTGG TGGCTTATGC 2700
 CTGTAATCCT AGAATCTTGG GAGGCTTAGA TGGGTGATC ACTTGAGCTC AGGAGTTCCA 2760
 GACCAGCCCG GGCTATGGC AAAACTCCGT CTCTACAAA AATAGAAAAA ATTAGCCAG 2820

WO 02/086443

PCT/US02/12476

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

TGTTGGTGGT	CATGCCCTGTA	GTCACAGTTA	CACGGCAGGC	TGAGGTGGGA	GGATCACTTG	2880
AACCCACAGG	GTCAGAGCTG	CAGTGAGCTG	AGATCACACC	ACTGTACTCC	AGCCTGGGTG	2940
ACAAAGTGAG	ACTCTATCTC	AAAAAGAAAT	TAGGATCAAT	TTGTCAATTT	CTACACAAC	3000
AACAACAAA	ACCCCTGTG	GGCACCTTGA	TTGAGATTGC	ATTGAATTTA	TATAAACTG	3060
TTGGGAGAAT	TGACATCTTA	ATAATATTGA	GTCCTCTGGC	CTATAAACAA	GGTCTGTCTT	3120
CCTAGGTATT	AATGTTTTGT	CTTCTATTTC	TCTTAATAAT	CTTTTGTAGT	TTTCAGTGTG	3180
CAGGTCTACC	ATGTCAGCAT	TTCATAGTTT	TGATGCTAAA	TGGTATTTTA	AAATTTCAAA	3240
TTCTAACAC	TTGTTGCTAG	TAAATAGAAA	TACAATTGAT	GTTGAACCTG	TATCCTTCAG	3300
CTTGCTTAAA	CTGTGAGTTC	TCATGGTGT	TTTGTAATTT	ACATCAACAG	TCATGTGTTC	3360
TATGAATAAA	GAGTTTTACT	CCTTC				

Seq ID NO: 156 Protein sequence:
Protein Accession #: Eos sequence

15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
MPCEKAMELY	RELHRAPEQ	LPANFEDGLR	QVLEEMKALY	EQNQSDVNEA	KSGGRSDLIP	60
TIKFRHCSLL	RNRCTVAYL	YDRLLRIRAL	RWEYGSVLPN	ALRFHMAAEE	MEWFNNYKRS	120
LATMRSLSGG	DEGLDITQDM	KPPKSLYIEA	GCSGAISAQ	ATSTSQVHLN	CNLHLPGPVS	180
KRLWRI						

Seq ID NO: 157 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 148-621

25
30
35
40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
TTCCGGCCCA	AAGCGCGGAG	CGGAGGCCGA	GGCGAGAGCC	TGGCGCTGTA	GGACTAGAAC	60
GAAAGAGTGT	AGGCGCCGAG	AGCCACAGATA	CCATTTTGGC	GTGAGAGCTG	GTGGTTGGCA	120
AGGCGCGGG	AGTGGGAAGC	GTCCGCCATG	TTCTGCGAAA	AAGCCATGGA	ACTGATCCGC	180
GAGCTGCATC	GCGCGCCCGA	AGGGCAACTG	CCTGCCTTCA	ACGAGGATGG	ACTCAGACAA	240
GTTCTGGAGG	AGATGAAAGC	TTTGTATGAA	CAAAACCACT	CTGATGTGAA	TGAAGCAAAG	300
TCAGGTGGAC	GAAGTGATTT	GATACCAACT	ATCAAAATTC	GACACTGTTC	TCTGTTAAGA	360
AATCGACGCT	GCACCTGTAGC	ATACCTGTAT	GACCGCTTGC	TTCCGATCAG	AGCACTCAGA	420
TGGGAATATG	GTAGCGTCTT	GCCAAATGCA	TTACGATTTC	ACATGGCTGC	TGAAGAAAGT	480
CGGTGTCTAA	AAGACTATGG	AGAATTGAA	GTTGATGATG	GCACCTCAGT	CCTATTAAAA	540
AAAAATAGCC	AGCACTTTTT	ACCTCGATGG	AAATGTGAGC	AGCTGATCAG	ACAAGGAGTC	600
CTGGAGCACA	TCCGTGCATG	ACCATGCGCC	GAGGCACCTC	CAGGCTTCAC	TCAACTCATG	660
GACTCCTCTG	TACTCACTCT	CTCCACCACCT	CCCTTCACCT	CCCTCTTTGA	TTTTAGAAGC	720
TATAGACATT	GTTTAAGATA	ACTAAGAATA	CTTGGCTAAG	AAGTATAAAT	TGCTAACTAT	780
TAAGGACTTT	CTTTTPTTAA	TGTTGTACAC	TATTTCTTCT	ACTCTTTTTT	GGTTTGGTTT	840
TTGTTTTGTA	GAGACTGTCT	CACATATGTT	CCCAAGCTGG	TCTCAAACTC	CTGGCCTCAA	900
GCAGTCTCTC	CACCTTAGCT	TCTCAAAGTG	TTGAGATCAC	AGGCGTGAGC	CACCTGCACCC	960
GGCCCCACT	CTTTTPTCTA	ATAAGCTGTA	TCTGTAAATCA	CAGCATTCCT	ACAGTTGTTA	1020
CAGTGTATGT	TTTAAATGAA	AGTAAACATG	GTTACATTTG	AATCTCTTAA	ATAAGCAGTC	1080
ACTTGCTCGG	ACAGGAAGAA	GGTAGATCCT	GTGTGTCTTG	TTTTCTGGTC	ATGTGTATTG	1140
TACAAGCTAG	AGAGCTGAAT	TTCTGAGATA	CACATTTTCA	AATCACATGC	AAGTGAAGAT	1200
GATGGTCTGT	AGAAATTTTC	AGTATATATA	ATGTTTAATG	ACATACTAAT	TTATCATCTG	1260
GCTATTGGG	AAGGAAGGAC	ACACATGGAT	TTTGACATTT	TCCACCATGG	TGGCTGGTGT	1320
GGCTTGTGGC	TATGGGGTGA	TCACCAATAT	CACCACTTTG	GAAAGGGACA	GTGAAATTTG	1380
GGCTAGAGAA	GGAACTTTTG	ACAGTTTTCC	CTGAGATTCA	GATTGACTGA	AAAGTCACAT	1440
GAAGAGTTGA	TTGCTTTTTT	ATGGTATGTT	TTAAACAGCT	GACATTTTAA	ATTTTGATGA	1500
AATCCAGTTT	ATTGCTTTGT	TCTTTTATGC	TTTGGGTGTT	GCATCCGAGA	AATCTTTTCC	1560
CATCCCAAGA	TCACAAATTT	TTTTCTTTT	TACTTCTAGA	AGTGTATATA	TTTTAAGCCT	1620
TATACTTTGG	TCTATGACCC	GTTTTTTTTT	TTGTTTTGTT	TTGTTTTTTC	GTGTTTTTCT	1680
TTGTTTTGAG	ATGGAGTCTT	GTTCTGTCTC	CCAGGCTGGG	GTGAGTGGC	GTGATCTTGG	1740
CTCACTGCAA	TCTCTATCCC	CTGGGTTCAA	GTGATCTCTC	TGTCTCAGCC	TCCCAAGTAG	1800
CTGGGATTAC	AGGCACAGGC	CGCCACGCCT	GGCTAATTTT	TGTATTTTAA	GTAGAGACAG	1860
AGTTTTACCA	TGTTGGCCAG	GCTGGTTTCA	AATCCTTGAC	CTCAAGTGAC	CCACCTTGCC	1920
CTCCCAAGT	TTTGGGATTA	CAAGTGTGGG	CCACCGCGGC	CAGCCTATGA	TCCATTTTGA	1980
ATGAATTTT	TATATGGTGC	AAGGTGTCAA	TCCACCTTCA	CTTTTCTTGG	GGAAATATAGA	2040
TATCCAGCTG	TTTCACTACC	ATTTTTTGAA	AGGACTGCCC	TTTGCTCTAT	CACCTTTTGA	2100
TTTTTGTAA	AAAGTAGTTG	TCAATGTATA	TGTGGGTTTA	TTTCAGGACT	CTGTTTTGTT	2160
CCATTGACCT	GTTTTTCTCT	CCTGAATGCC	AATACCATAT	TTGTATGTAG	TGTATGTAA	2220
TTTCTAATAA	TTCTTGAAAC	AGATAGTATT	AATGTGTGAT	ATTTTTGCTG	TTGTTGTAT	2280
TTTTTGTAGA	GATGGGTTCT	CACCGTGTG	GCCAGGCTGT	GTTGAACTCC	TGAGCTAAAG	2340
CAATACACTT	GCTTGTCTCT	CCCCATGTGC	TGGGATTACA	GGCGTGAGCC	TTGGTGCTGG	2400
CCCACTGTAC	CACATTTCTT	TTTGAGATT	GTTTTGGCTA	TGTTAAGTCC	TTTGCTTTTG	2460
ATGTGAAAT	TGGGAACAGG	CAGGTTGTGG	TGGCTTATGC	CTGTAATCTC	AGAACTTTGG	2520
GAGGCTTAGA	TGGGTGGATC	ACTTGAGCTC	AGGAGTTCCA	GACCAAGCCC	GGCCTATGGC	2580
AAAACCTCCG	CTCTACAAAA	AATAGAAAAA	ATTAGCCAGG	TGTGGTGGTG	CATGCCCTGA	2640
GTCACAGTTA	CACGCGAGGC	TGAGGTGGGA	GGATCACTTG	AACCCAGAG	GTCAAGACTG	2700
CAGTGAGCTG	AGATCACACC	ACTGTACTCC	AGCCTGGGTG	ACAAAGTGAG	ACTCTATCTC	2760
AAAAAGAAAT	TAGGATCAAT	TTGTCAATTT	CTACAACAAC	AACAACAAA	ACCCCTGTTG	2820
GGCACTTGA	TTGAGATTGC	ATTGAATTTA	TATAAACTG	TTGGGAGAA	TGACATCTTA	2880
ATAATATTGA	GTCCTCTGGC	CTATAACAA	GGTCTGTCTT	CCTAGGTATT	AATGTTTTGT	2940
CTTCTATTTC	TCTTAATAAT	CTTTTGTAGT	TTTCAGTGA	CAGGTCTACC	ATGTCAGCAT	3000
TTCATAGTTT	TGATGCTAAA	TGGTATTTTA	AAATTTCAAA	TTCTAACCC	TTGTTGCTAG	3060
TAAATAGAAA	TACAATTGAT	GTTGAACTTG	TATCCTTCAG	CCTTGCTAAA	CTGTGAGTTC	3120
TCATGGTGT	TTTGTAATTT	ACATCAACAG	TCATGTGTTC	TATGAATAAA	GAGTTTTACT	3180
CCTTC						

Seq ID NO: 158 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51

MFCEKANELI RELHRAPEGO LPAFNEEDGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIP 60
TIKFRHCSLL RNRRTCTVAYL YDRLLLRIRAL RWEYGSVLPN ALRFHMAEE VRCLDYGEF 120
EVDGTSVLIL KQNSQHLPLR WKCEQLIRQG VLEHILS

Seq ID NO: 159 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 149-229

1 11 21 31 41 51
GTTGCGCGCC AAAGCGCGGA GCGGAGGCC AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60
CGAAAGGAGT GAGGCGCGGA GAGCCAGAT ACCATTITGG CGTGAGAGCT GGTGGTTGGC 120
AAGGCCGCGG GAGTGGGAAG CGTCCGCCAT GTTCTGCGAA AAAGCCATGG AACTGATCCG 180
CGAGCTGCAT CGCGCGCCCG AAGGGCAACT GCCTGCCTTC AACAAATTAG TGGGTGTGGT 240
GGCACACACC TGTAATCCCA GCAACTTAGG AGGCTGAAGT GAGAGGATTG CATGGCTCCA 300
GGAAGTTGAA ACTGCAGTGA ACTGTGGTCA CGCTATTACA CTCCAGCCTG GGTGACAGAC 360
TGAATCCCTG TCTCAAAAAG GAAAGGAGG ATGGACTCAG ACAAGTCTG GAGGAGATGA 420
AAGCTTTGTA TGAACAAAAC CAGTCTGATG TGTCTCTGT TAAGAAATCG ACCTGCACT 480
GTAGCATACC TGATATACCG CTGCTTCGG ATCAGAGCAC TCAGATGG

Seq ID NO: 160 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
ATGTTCTGCG AAAAAGCCAT GGAAGTATC CGCGAGCTGC ATGCGCGGCC CGAAGGGCAA 60
CTGCCTGCCT TCAACAATTA G

Seq ID NO: 161 DNA sequence
Nucleic Acid Accession #: U10694
Coding sequence: 1333-2280

1 11 21 31 41 51
GGATCCGGCC GGATCTCAGG GAGGTGAGGA CTTTGTCTC AGAGGGTGTG TGTGGACAAA 60
ACAGGGAGGC CCTGTGTTCC ACAGACACAG TGGTCCCGAG ATTGGAGAGC AGTCCAGGTG 120
AGGAACCTAA GGGAGGATCG AGGGTAACCT CAGGCCAGAG AAATCTCTCAG ATCAAGAGAG 180
TTTGCCCTGC CCTACTGTGC ACCCCAGAGA GCCCGGGCAG GGCTGTCTGC TGAGGTCCCT 240
CCTTTATCTT GGCATCACTG GTGTGCGGGA GGGCTGGCCT TGGTCTGAGG GGGCTGCCT 300
CAGCTCAGCA GAGGAGGGGT CCCAGGCCCT GCCAGGAGTC CAGGTGCAGA CTGAGGGGAC 360
CCCCTCACC AAACACAGAG GACCTAGCCC CACCCTGCC CTGTGTCTAG CTGAGGGAG 420
CCGCTGGGTG GATGAGCTCC CCTCACTTCC TCTTCAGGTG TCTCCTGGAG ATAGGGCCTC 480
AGGTCAACAG AGGGAGGGGT CCAGACCTTG CAGGCATCAA GATGAGGACC AGGCAGTATC 540
CTACCCAGG GACACATGGA CCCCATGAA TTTAGACATC TCTTACTGTA CTTCCAGGA 600
AACCTTGGC AGGTGTGGGC AGATGTTGGT TGGGGCATGT CTTCTGTTC CATATCAGGG 660
ATGTGAGCTC CTGATCTGAG AGACTCTCAG GCAAGTAGAG GAGTAGAGTC CAGTCCCTGC 720
CAGGAGAAAG GTGAGGGCCC TGAGTGAGCG CAGAGGGGAC CATCCACCCC AAAAGTGTGT 780
AGAACTCAAG AGTGTCCAGC CCGCCCTCTT GACAGCACTG AGGGACCGGG GCTCTGCCCTG 840
CAGTCTGAG CCTAAGGGCC CCTCGATTCC TCTTCCAGGA GCTCCAGGAA GCAGGGCAGGC 900
CTTGGTCTGA GACAGTGTCC TCAGGTGCGA GAGCAGAGGA GACCCAGGCA GTGTGAGCAG 960
TGAAGGTGAA GTGTTCAACC TGAATGTGCA CCAAGGGCCC CACCTGCCCC AGCACACATG 1020
GGACCCCTAA GCACCTGGCC CCATTCCCTT TACTGTCACT CATAGAGCCT TGATCTCTGC 1080
AGGCTAGCTG CAGCTGAGT AGCCCTCTCA CTTCTCCCT CAGTGTCTCG GCACAGGCTA 1140
ACCAAGGAGA CAGGAGCCCC AAGAGGCCCC AGAGCAGCAC TGAAGAGAC CTGTAAGTCA 1200
GCCTTTGTTA GAACCTCCAA GGTTCGGTTC TCAGCTGAAG TCTCTCACAC ACTCCCTCTC 1260
TCCCCAGGCC TGTGGTCTCT CATGCCCCAG CTCCTGCCCA CGCTCTGAC TGCTGCCCTG 1320
ACCAAGATCA TCATGTCTCT CGAGCAGAGG AGTCCGCACT GCAAGCCTGA TGAAGACCTT 1380
GAAGCCCAAG GAGAGGACTT GGGCTGTATG GGTGCACAGG AACCCACAGG CGAGGAGGAG 1440
GAGACTACCT CCTCTCTGA CAGCAAGGAG GAGGAGGTGT CTGCTGCTGG GTCATCAAGT 1500
CCTCCCAAG TCTCTCAGG AGGCGCTTCC TCCTCCATT CCGTCTACTA CACTTTATGG 1560
AGCCATTTCG ATGAGGGCTC CAGCAGTCAA GAAGAGGAAG AGCCAAGCTC CTCGGTGCAG 1620
CCAGCTCAGC TGGAGTTCAT GTTCCAAGAA GCACGTGAAT TGAAGTGGC TGAGTTGGTT 1680
CATTTCTGCT TCCACAAATA TCGAGTCAAG GAGCCGGTCA CAAAGGCAGA AATGCTGGAG 1740
AGGCTCATCA AAAATTACAA GCGCTACTTT CCTGTGATCT TCGGCAAGGC CTCGAGTTC 1800
ATGCAGGTGA TCTTTGGCAC TGATGTGAAG GAGGTGGACC CCGCCGGCCA CTCCTACATC 1860
CTTGTCACTG CTCTGGCCCT CTCGTGCGAT AGCATGCTGG GTGATGGTCA TAGCATGCC 1920
AAGGCCGCC TCTGTATCAT TGTCTGGGT GTGATCTTAA CCAAGACAA CTGCGCCCT 1980
GAAGAGGTGA TCTGGGAAGC GTTGAGTGTG ATGGGGTGT ATGTTGGGAA GGAGCACATG 2040
TTCTACGGGG AGCCAGGAA GCTGCTCACC CAAGATTGGG TGCAGGAAA CTACCTGGAG 2100
TACCGGCAGG TGCCCGCAG TGATCCTGCG CACTACGAGT TCCTGTGGGG TTCCAAGGCC 2160
CACGCTGAAA CCAGCTATGA GAAGTCTATA AATTATTGG TCAATGCTCA TGAAGAGAG 2220
CCCATCTGCT ACCCATCCCT TTATGAAGAG GTTTTGGGAG AGGAGCAAGA GGGAGTCTGA 2280
GCACCAAGCG CAGCGGGGCG CAAAGTTTGT GGGGTGAGG CCCCATCCAG CAGCTGCCCT 2340
GCCCATGTG ACATGAGGCC CATTTCTGCG TCTGTGTTG AAGAGAGCAA TCAGTGTCT 2400
CAGTGGCAGT GGGTGGAGT GAGCACACTG TATGTCACT CTGGTTCTCT TGTCTATTGG 2460
GTGATTGGGA GATTATCTCT TGCTCCCTTT TGAATTGTT CAAATGTTCT TTTAATGGTC 2520
AGTTTAAATGA ACTTCACCAT CGAAGTTAAT GAATGACAGT AGTCACACAT ATTGCTGTT 2580
ATGTTATTGA GAGTGAAGAT TCTTGTCTTT GAGTCACATG GGGAAATCCC TGTATTTTG 2640
TGAATGGGA CAGATAACA TAGCAGAGGA ATTAATAATT TTTTGAAC TTGAACCTAG 2700
CAGCAAAATG GAGCTCTAAA AGAAATAGT AAATGAAAT GTAGTTAAT CTGCTTAT 2760
ACCTCTTCT CTCTCTGTA AAATTAAC ATATACATGT ATACCTGGAT TTGCTGGCT 2820
TCTTTGAGCA TGAAGAGAA ATAAAAATG AAAGAATAAT TTTCTGTG CACTGGCTCA 2880
TTTTTCTTC AGACACGAC TGAACATCTG TTATTCCGAA CACCTGGGT T

Seq ID NO: 162 Protein sequence:
Protein Accession #: AAA68877.1

1 11 21 31 41 51
MSLEQSRSPHC KPDEDLAQQ EDLGLMGAQS PTGEEBETTS SSDSKKEEVS AAGSSSPPOS 60
PQGASSSSIS VYITLWSQFD EGSSSQEEEE PSSSVDPQL EFMFQELKL KVDELVHLL 120
HKRYVKEPVT KAEMLSEVVK NYKRYFPVIF GKASEPMQVI PGTDVKEVDP AGHSYILVLA 180
LGLSCDSMLG DGHSMFKAL LIIVLGVLIT KDNCAPEVI WEALSVMGVY VGKEHMFYGE 240
PRKLLTQDWV QENYLEYRQV PGSDPARYEP LWGSKAHEST SYEKVINYL VMLNAREPICY 300
PSLYEEVLGE EQEGV

Seq ID NO: 163 DNA sequence
Nucleic Acid Accession #: AF292100
Coding sequence: 30-809

1 11 21 31 41 51
GGGGGGGAG AGGCCTGGAG GACACCAACA TGAACAAGTT GAAATCATCG CAGAAGGATA 60
AAGTTGCTCA GTTATGATC TTCACACAA CTAGTGAAAA AACAGCAGTA AGTTGTCTTT 120
CTCAAAATGA CTGGAAGTTA GATGTGCGAA CAGATAATTT TTTCAAAAT CTGGAACCTT 180
ATATACGAGA GAGTGTAAAA GSATCATTGG ACAGGAAGAA GTTAGAACAG CTGTACAATA 240
GATACAAAGA CCTCAAGAT GAGAAATAAA TTGGAATAGA TGCCATACAG CAGTTCTGTG 300
ATGACCTGGC ACTCGATCCA GCCAGCATTA GTGTGTTGAT TATTGCGTGG AAGTTCAGAG 360
CAGCAACACA GTGCGAGTTC TCCAAACAGG AGTTTCATGGA TGGCATGACA GAATTAGGAT 420
GTGACAGCAT AGAACAACTA AAGGCCCGA TACCCAGAT GGAACAAGAA TTGAAAGAAC 480
CAGGAGGATT TAAGGATTTT TACCAGTTTA CTTTAAATTT TGCAAGAAT CCAGGACAAA 540
AAGGATTAGA TCTAGAAATG GCCATTGCCT ACTGGAAGTT AGTGCTTAAT GGAAGATTTA 600
AATTCCTAGA CTTATGGAAT AAATTTTGTG TGAACATCA TAAAGCATCA ATACCAAAAG 660
ACACTTGGAA TCTCTTTTA GACTTCAGTA CGATGATTGC AGATGACATG TCTAATTATG 720
ATGAAGAAGG AGCATGGCCT GTTCTTATG ATGACTTTGT GGAATTGCA CGCCCTCAAA 780
TTGCTGGGAC AAAAGTACAA ACAGTGTAGC ACTAAAGGAA CCTTTTAGAA TGTACATAGT 840
CTGTACATAA AATACAAACG AAAATTGCAC AGTCAATTTG TGCTGGCTGG ACTGAACTGA 900
AGATCAATCC TCACAATTCA GACTGAGGCT TGACACAAA CTTTAAGGAT ACATCTTGGG 960
CCATATCGTA TTTCATTCTT CTAAATGGTG TTTGGGCTTG TCTTCTAGTC TGGGCCGCTC 1020
TAAACATTTA TAATCCAAAC ATTTGTGATG TCATCTTATA TCTGTGGACC ATCCTAGTTT 1080
ATTCTCCCAT AAGTCTTAGA AGCTTTATGG TGATTATTTT GAGGTTTCA TTCTGCGATA 1140
AAGCACAATG CTGTCTTCAT CAGAAAACAG TTGGCATAAG AATTAACAT ATGAACATCA 1200
CAAAACAATT TATAAAACTT TCTTAAATAT ACGCTTTGGG CTAGTTGCAA AGACTATGCT 1260
AATAGCATT CCAGTGAGAG TGATATATTT AAGTGTACTG GATCTGGAAT GGTGTTTTGG 1320
TTTGGGGGGA ATTTTTTTTT TTTCCTGGCA AATCACATAT GTTGTGATG TGAGTATCTG 1380
ATGAAGAAAC AATGTACAGAA TAACCGACAT GAAAATTTT TAGGATAACT TGGTGCCTAC 1440
CTGAAAAATG TATTGTGTTT TAGACTCTTG ATTTCAAAAG GTTCCACAGA ACTAGTCTGC 1500
GCTTACCTTA CCCATGTTTA TATATAGCTG TCCTACAGGG AGCTTTTATT TAGAAAATGT 1560
CTGCATAATG TTAGATTCTT CTCTGTCTCA CATTATGCAC TACATAATTG GACTTCATTA 1620
TGCTTTTGAA ATGCTTATCT GCCTGTGACA TAAGTTAAAC TATTTAATTT GTTTTGAATG 1680
TTTTGGATG CTACACAATA CAATATTCTA AATTTAGGCA TGAGGGTTTT TTTGTTTTAT 1740
TTTTACTTTT TTTTGTTCAT TGCACTATGG AACACAAATG AATTTCTCTT AATTTATAAG 1800
AAGATAGTAG GAGTTAAATT TTGAAAATGG TTGTGATGAG CCACGAAAT CAATCTTTAT 1860
AATATAGGTA CTGCTCTTTC AGACAAACAG TCCATTTTAA ATGACTTCTT ATTTTGTGTA 1920
AATTACTTTA ACTGCTAATC ACTGTGGTGG CCAAAATTTT ACTTCAGAAG CAAAGATTTT 1980
CAACAAAGCA AATGCTATGC AAAATACCA TCTGGCTTCT AGTCTATTTA CTGTTTGTGT 2040
TCACTCAGAT TAGCTCAGTT TTCTCATCAA AGCAGAAATG TATCTTGCCT GTGTGTGTGT 2100
GTGTGTGTGT GTGTGTGTGT GTATGTGTGT ATATATATAT ATATATATAT ATATATATTT 2160
TTTTTTTTTT TTTTTTTTAA ATTACAAAAG CCATGAGCTG CTTTATGCTG GAAAATGCTC 2220
ATTTCCCTGT TCACTTACTG ACATGTGAAG AAGGGTTTCT TGCTTTCTTA AACATTCCG 2280
TAAGCGAGCG TAGAATAGTA ATACTTCAAA TGTTTGATGA TTATGCTCTT TTGATAGGAA 2340
TAGATTCTGC TTGGGATATA TATCCAGGCA CTCTCTAAGG TCTAGGGTTG ATATTAACAA 2400
AGGAATGTAC TTGAATAGC AGTACATTTT ATGCAAAATG GGAATATTAT TTAAGAAACA 2460
ATGACATATC AAACTGCTTT TTTACATGAT TTTGAAATAG ACTAGAAAGC TTTCCCTATA 2520
GACATATTAA TATTCATCT ATAACTTTAA TTCAAGAAATG CAGTTTACC AAAAGAAAAA 2580
TTTGAATAAT TCTATTACGG CTACTGGAAT TGGTTATTAA AAGAAAAAGG AAAAGAAAGA 2640
ATCTTGCTGC TTTCAATATT TCCTGATTTT TTTGTAATAA TAAAGAGGAA CTTCAATTAT 2700
GAAAAATTTT TAAAGATAT ATATATCTAT ATATCTATAT ATATGACTG TTTTGTTCCT 2760
TGCTTGAAAG ATTTTGAGTT ATGTTTATG GTTTCAGATT GATTAAATCA CATATGCTGT 2820
GTTTCTTTTA AAAGTCATAT GGGTTCGTGG CCTAATGCCT TGGATTTTAC ATATTTTCT 2880
TTTTAAATGC AAAACCTTTT CAACAAAAATA GTGTTTGTCA TCAGGTTGGT ACTAAACATT 2940
TATAATTACT GTGTAATTAT AAACAAAAAT ACATAAAGCT TTGAATATAA TTATGTAGCA 3000
TAAAGTTTAA GGTGTGTCAC TATGATGGCA TCTTGAATTT AAACAAAAAT TTTACTAGGG 3060
CTGAAAAGAG AAGACTGATT TAATGTGGTG TGATTATTCT GAAGATAAAT GTCTGGCTAC 3120
AGGGAATATT TTGTACTAAA AAATGATTAC ACATATGGCT GTGTGTGTTT GAGTCTGTGT 3180
CTGTGAGAGA GCCAGAGAGA GTGAGAGAGA TTGACAGAGA AAGGAGAGA CACACACAG 3240
CCCTTGAAAT TGCTTAACT CCTAAGTGT TCACTCCTCA TTCCGGTAAA CTCCCCATGC 3300
TGATCTTTG TTTTAACTG AACCATAGGT ACAGTTTCTT TTTTGCCAAA TGTCAAAACA 3360
GGTACAAATT TTAATATGTA ATGCTTTTAA AATAGAAAAA TGTATAAAAT TAGAAGTGCC 3420
CATATATAAA AATACTTTGA GATGAAGATT ATCTTTAGTG AATATCATCT GCATATCTCT 3480
GTAAGTTCAA TTGTGTTTCT TACAGTCCCT GTCATATTAC CAACAGAGGC AATAAAAGCT 3540
GCAGTGAAT TG

Seq ID NO: 164 Protein sequence:
Protein Accession #: AAG00606

1 11 21 31 41 51
MNKLKSSQKD KVRQFMIFTQ SSEKTAVSCL SQNDWKLDVA TDNFFQNPFL YIRESVKGSL 60
DRKLEQLYN RYKDPQDENK IGIDGIQDFC DDLALDPASI SVLLIAWKFR AATQCEFSKQ 120
EFMDGMEBLG CDSIEQLKAO IPKMEQELKE PGRFKDFYQF TTFNAKNPGQ KGLDLEMAIA 180
YWNVLNLRFP KFLDLWNKFL LEHKRSIPX DTWNLLDIFS TMIADMSNY DEEGAWPVL 240
DDFVEFARPQ IAGTKSTTV

Seq ID NO: 165 DNA sequence
Nucleic Acid Accession #: AF256215
Coding sequence: 220-2028

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
CTCCAGTCCG CATGCTCAGT AGCTGCTGCC GCGCCGGCTG CGGGGCGGCG TCCGCTGGCG 60
GCTTACGGGC TGCGGTGGGG GCGGCGCGGG CACCGCGCAG GCGCGGCCAG TCCCGGCTTC 120
CCTGCTCCAG AGCCGCGCGC TGCGCGGGGG CAGGCGGGGC CCGGGGCTCC TCCATGCTGC 180
CAGCGCGCGG GCTCGGGAGC CGACCAAGTG GCTCCTGCGA TGGCGGCGGA AGAGGAGGCT 240
GCGGCGGGAG GTAAAGTGTT GAGAGAGGAG AACCAAGTCA TTGCTCCTGT GGTTCCTCAGC 300
CGCGTGAGTC CAGGAGACAG ACCAACAGCT ATGGGCTCTT TCAGCTCACA CATGACAGAG 360
TTTCCACGAA AACGCAAGG AAGTGATTCA GACCCATCCC AAGTGAAGA TGGTGAACAC 420
CAAGTTAAAA TGAGAGGCTT CAGAGAAGCT CATAGCCAAA CTGAAAAGCG GAGGAGAGAT 480
AAAATGAATA ACCTGATTGA AGAACTGTCT GCAATGATCC CTCAGTGCAA CCCCATGGCG 540
CGTAAACTGG ACAAACCTTC AGTTTAAAGA ATGGCTGTTC AACACTTGAG ATCTTTAAAA 600
GGCTTGACAA ATTCTTATGT GCGAAGTAAT TATAGACCAT CATTTCTTCA GGATAATGAG 660
CTCAGACATT TAACTCTTAA GACTGCGAGAA GGCCTCTTAT TTGTGGTTGG ATGTGAAAGA 720
GGAAAAATTG TCTTCGTTTC TAAGTCAGTC TCCAAAATAC TTAATTATGA TCAGGCTAGT 780
TTGACTGGAC AAAGCTTATT TGACTTCTTA CATCCAAAAG ATGTTGCCAA AGTAAAGGAA 840
CAACTTCTTT CTTTGTATAT TTCAACCAAG GAAAGCTAA TAGATGCCAA AACTGGTTTG 900
CAAGTTCACA GTAATCTCCA CGCTGGAAGG ACACGTGTGT ATTCTGGCTC AAGACGATCT 960
TTTTCTGCTG GGATAAGAGG TTGTAATAAT TCTGTCAAAG AAGAGCATGG ATGCTTACCC 1020
AACTCAAATT CAGTAAGAGC CAGAAAATTC TATACATACC ATTGCATGG TACTTGGAGA 1080
AGCTGGCCTC CAAATATTGT TGGAAATGAA GAAGAAAGGA ACAGTAAGAA AGACAACAGT 1140
AATTTTACCT GCGTTGTGGC CATTGGAAGA TTACAGCCAT ATATTGTTC ACACAACAGT 1200
GGAGAGATTA ATGTGAAACC AACTGAATTT ATAACCGGCT TTGCACTGAA TGGAAAAATT 1260
GTCTATGTAG ATCAAAGGCG AACAGCGATT TTAGGATATC TGCCCTCAGGA ACTTTTGGGA 1320
ACTTCTGTGT TCTTCGTTTC TCATCAAGAT GACCACAATA ATTTGACTGG CAAGCACAAA 1380
GCAGTTCTAC AGAGTAAGGA GAAATACCT ACAGATTCTC ACAAATTCAG AGCAAAAGAT 1440
GGCTCTTTTG TAACCTTAAA AAGCCAATGG TTAGTTTCA CAAATCCTGT GACAAAAGAA 1500
CTGGAATATA TTGTATCTGT CAACACTTTA GTTTTGGGAC ATAGTGAGCC TGGAGAAGCA 1560
TCATTTTAC CTTGTAGCTC TCAATCATCA GAAGAATCCT CTAGACAGTC CTGTATGAGT 1620
GTACTGGGAA TGTCTACTGG AACGTAAGT GGTGCTGGTA GTATTGGAAC AGATATTGCA 1680
AATGAATTC TGGATTTACA GAGGTTACAG TCTTCTCAT ACCTTGATGA TTGAGTCCA 1740
ACAGGTTTAA TGAAGATACA TCATCTGTA AACTGCAGGA GTATGTCAA TAAGGAGTTG 1800
TTTCCACCAA GTCCTCTCTA AATGGGGGAG CTAGAGGCTA CCAGGCAAAA CCAGAGTACT 1860
GTTGCTGTCC ACAGCCATGA GCCACTCCTC AGTGATGGTG CACAGTTGGA TTTCGATGCC 1920
CTATGTGACA ATGATGACAC AGCCATGGCT GCATTATAGA ATTACTAGA AGCAGAGGGG 1980
GGCCTGGGAG ACCCTGGGGA CTTCACTGAC ATCCAGTGA CCCTCTAGCC TTTGATTTT 2040
AACTCCAAA ATGAGAAAAC TTTTAAAGCA TTATTACGA AAAAAGCTGC TCAACTATTC 2100
TTAAGTACTG TATTGATATT GTTTGTATCT TTTATTAATG TTCTACCCTT TTTATAGAT 2160
TTGCACTCTC TGCTCAGCAG GATGTGGGGA AATACGTTTT CTTCCCAAGA GAACCAAGTT 2220
TATTATAGAC CTCTTATCTC AGTGAATAGG CTTATAATCC ACTAGTTGCC ATATTTTTCG 2280
TAAATATTTT CTAACCAAGA ATACTACTTA CATATTGTTT TGGCTTTGTT TTAATTTTGA 2340
TGCAGTTTTT TTTAGTTGAG GTAATGTAAT ATATTGATGT TTTCTTTTGT GTCTAAGATT 2400
GATTATAAT AGTAGGTTTG TATAATTTGG AACATTTTCC ATGCCITGGG AATTTCTCTA 2460
ATTGAGGATA GGGCTTACAC ACTTTAAGAA AACAGTGAGT ACTTGAACAT TTAAGGGGAC 2520
AGTCAATTTT ATAGTCATA TCACATTGAA TACTGTATTT GATCTTTGGA GACTTAGGCA 2580
AGCAGAGAGT TGGGATATTT ATGCTCAGTT GAGCACTTTA AGATGAATTT TAAGTGAGAT 2640
GATTTCTGCG TTAAGACTCA GAAAGTCAAA AGAGTTTCAG CTTTCTTAC AGAAAAGGAA 2700
GGATCTTGGG CCCTAGATCT TGGGGATTAA CCTCTGCATA TAAGATTTAC TCTTAATAGG 2760
CCAGAGCTGG TGCTCAGCGG TGTAAATCCA GTACTTTGGG AGGCTGAGAG GGGCAGATCA 2820
CTTGAGGTACA GGAATTCAG ACCAGCCTGG CCAATATGGT GAAACCCCGT TTCTACTAAA 2880
AATACAAAAA AAATTAACCA GGCCTCACTT CTTGAGGTAA CTAACCAACT CCCACGATAA 2940
ACTACAGTCA TCTATGAGCG CAAAGGCTC ATGACCTAAT GGCACACACC TGTAAATCCA 3000
ACTGCTTGGG AGGCTGAGGC GAGAGGATTG CTTGAACCTG GGAGGCAGAG GTTGCACTGA 3060
GCGGAGATCG CACCACTGCA CTCAGCTCTG GGCAACAGAG TGAGACTTCA TCTCAAAAAA 3120
AGTAAAAAAA AAGATTATAT ATAATCACTG AAGATCTCTA TTATAGATAG ATTAGGTTTT 3180
TGACATTGGA AACATCTTA GGGATAGATT TGTCTTAAAG GAAAAAGTA GCGCCGGGCA 3240
GATTAATGTT CTTGTGTAAA GTCACACATT AAATTCAGTC ACACATTAAA TCTATAGAGT 3300
TTTAAATGTT TAATGTATAT AAACCACTTT CTTTATACAC ATTTGGGAAA ACATTTGTTT 3360
CACAGATTAA ATGATTAACT AACTGACCCA GGAAGTATT GTAGCTTTCT AAGTAATTAG 3420
GCAATTACAG TTATTGCTGT TAACCAAAGG TAATAAAACA AAATGACAAG TACATGTTTA 3480
AAATTATGAG GCAATGAGAA ATAATTAAA AACCAATTTT CTAGTTATAA TTTAAATTTT 3540
GGAGAGCATT TTTAAGCATA ATTAATCCAG AGGTGGCTCA AATTGAGTAT AAGAATTAA 3600
ATTATTAAAA ATACTGCATG TCTACCTTCT CGGGGATCAT ACTTTATAAC ACTTTCTGCT 3660
TCAGTAGCTC TTATAGCTTT GCCAAGTATG CTCCCATATT TTCTCTCTG TGCCCTGCAA 3720
ATGAAAGTCA GATAGGCTGG GAACTCATGG GGCAGCCCTC AGACTTCAAT GTGGGCTTCA 3780
AATCCAGTTT CCGTCTCTAT ATGGTGCTAC ATCTTCCAG AAAATTTCCC TCAGAGCCCC 3840
TCGCCAAATC AAAGCATTAT TTTGACCCCT CATGCTATTT CTTTAGCTGT AGGTGATAGA 3900
TTAGAACTTC TGTCAGACAT GTTAATGACA AACATACCAA CAGACAATAA CCAAAGCAAA 3960
TGTTCTCTTC AAGTGTGAAA TGTGCAAGGG CTCGTGGGCA AGGATGTATT GGCACACTGT 4020
CCTCTTGAAC TGATAGTGTG CCAGCAATGT TGGAGGTTGG CACCATCTCT GGTCCGACAC 4080
TTGAGGACCT GAGAGACATC AGGTTTAGAA TGAGCCAAAG AAATCTTACA AGATGGGGAG 4140
AATTTGTTGG CAGCAGCTTA AGTGTATAG TTAAGTCTAA AGAAGTATGA AAGATCCCTC 4200
GTGTCTCTTA AATTGAGCAG AGGGGCTGCT CTACCAATAT CACTTTTTAG GGGACTGAAC 4260
CATTGAGGTT TAGACTTGCG TTCCAAGAG TCTGCTTAAG CCAGGGGTGG CAGGGTAGGC 4320
CATCATAGCT GATAGGCTCT AAAAGCAGAT GGGGGCAGAC TTGCCCTCGT GATGCCAGGA 4380
TTTGAGAGGC AGAGTTTCTA GAGGGAGACC AGTGCTGCTC CTCACAGTGG CAGTTTTTTC 4440
TCTTTGCAAG AGGAGGGGCT GTTCAATTCC ATAGACCAGT GGGCAGATAG CCAAGTTGAAT 4500
ACTCTGTGCA TGGTTTGATC CTTTATFAGT TCGCTCTAAT ATTTTCTGT AGATCCTTTT 4560
GTCCCTGAGC CAAATCTTAA TCCATGCATT GTATGATACC GTAGCTCTCC TAAGGTTTGT 4620
GTTCTCTTCA AATGTCTTTA GTTTCTTCA ACTAAATTTG ATTTTGTGCT TTAGAAGTGA 4680
CATATTTTGA TGGTATACAC TATGTTCTCT TTTTCTACTG CGAGTCAATT TTTTGAATTT 4740
TCGTGAGAAA GAATATATCT ACAAAATGCA CGAAAGTATC ATAAAAACAG TACTCTAGAG 4800

5
10
15
20
25
30
35

CAGCGCTGTC CAATAGAAAT ATAATCTGAG CCACATGTAT AATTTTATTT TCTTCTAGCC 4860
ACATTAAAGA AGTAAAGA TAGAAGTAGA ACTAATTTTA ATGTTTAAAT TCAGTATATC 4920
CAAAATATCA TTGGAACATG TAATTAATAT AAAATTATTA ATGTGATATT TTACATTCTT 4980
TTGGTAATAC TAGTCTTCAA AATCTGGTAT GTATCTTACA TTGATAGCAC ATCTCACTTT 5040
GCTACTAGCA CATTGCAAGT GCTCAGTAGC CACATGTGGC TAGTGGCTAC TGCCTGGAC 5100
AGCACAGTTC TAGTTCCAC OCTAACACCC AAGTCTGTG GATTAGAAAT CCAGAAATCAG 5160
AGCTGGAAGT AAACATAGAG ATCAAACCTC CTTTAAAAA TGAGGACGCT GAGGCACAGA 5220
GTTTAAATGG CTTCGATGAG GTCATACAGC TAAATTGAGC CTCACAGGG TCTTCTGATT 5280
CCAGGCACTC TTCCCACTCC ACTACATTAC TGTAGTGGTA ATTCTTAGGG TTAATAAAG 5340
TGTAGATAG GCCGGGGCGA GTGGCTCATG CCTGTAAATCC CAGCACTTTG GGAGGCCGAA 5400
GTGGGGGAT CACGAGGTCA GGAGATCGAG ACCATCTGCG CCAACATGGT GAAACCCCGT 5460
CTCTACTGAA AATACAAAGC AAAATTAGCC AGGTGTGGTG GCGGGCGCCT GTGGTCCAG 5520
CTGCTCTGGA GGCTGAGGCA GAATGGCGTG AACCCAGGAG GCAGAGATGG CAGTGAGCCA 5580
AGATCGCGCC ACTGCACCCC AGCCTGGGCG ACAGAGCGAG ACTCCATCTC AAAAAAAG 5640
AAAAAAGAA AAGAAAAAG AAGAAAAAGT TAGAGAACAT TATATTAAGT GGTATTATT 5700
GAAGTAGACC AAAGTTTATA CCATAAGGAT ATTTTCTCTT AAATACCATG TTGAAGAAG 5760
AATTATTAT TGATCCTTGA ATCTGTAAGA TCAATAACA AGTCTCTATC CATGTTACCA 5820
AATTAACTT TTGAATAA ATAACTTTA AATATCAGA TGTGTTATTA CAGGATGATA 5880
CTTGAATCA AGTGAATGA GTTATATGGT CATCACTAAA TTTAGAAATC TATTGTGAA 5940
CAAGACAAA CAGGAAGATA CAGAATAGAG ACTTTTAGTA AATAAATGGA ATTTAAAGA 6000
AAGGTGTTAT TTACAGTGTG ACGACAGAAA AGGATGTCTT TGTGTGATA GTCTTTGAGG 6060
GATCTCGTA AATCTGGG CACAGGTACA AGAATAGCC AATATTAGT TCCAGACCA 6120
TGTATTAGT TGTCCAGTTT CAGATCATGC TGCCAGAGG TATCTCCCC TCAGGTGGGT 6180
CATCACTGAG CCGTGGAAAT GGAGACTCAT ACTTGGCCAG CACAATGTTA CGGCGACACA 6240
GGCGACATC TATGATTAGC TAGAAGCCAT AAAGAAAGC TGCTAAGTGG CCACTAGGTG 6300
CCACTTTCT GTTTTGTAA TGCTTTCAIT AGCAGATCTT TTTTTCCTCA GCTCCATGGG 6360
GCCTATGAGA GGCATTATG ATTTTGTGC CTACAATAAG TCAGCCTGTC TGGTGTGAGT 6420
TGTTTTATGA GAAATGCTTT CCAAGGGAGG TCTAGGAAGA TCCTGACACA TAAGAACTTT 6480
GGCTTAGAGA GCTTTCCAGG TGTAGTGCCA ATAAAACTG ACCTGGAAAG AAAACCTGCC 6540
CAGCAGGAA CATGCTTTCT GAACTCACTT GAGAGTGTAT GGTGTATGTC ACTTCTCATA 6600
TATCTGTAG TTTAGATTG TCTTTTATAC AATTTTATAG TCTTTTCCAG TTCACPTGTG 6660
CTCGTCTGTA TTTGGTATT TTTAAATTT TGTGGTAAAT AATGAAAGA GTGAATTTAT 6720
ATTTTATAAT TACTCATTTG TAGTTTTTTT TTTTAAATTA ATAACTTCC TCCAAAAAGT 6780
GCTCCCTTAA AA

Seq ID NO: 166 Protein sequence:
Protein Accession #: AAG34652

40
45
50

1 11 21 31 41 51
MAAEERAAAG KGLVREENQC IAPVSSSRVS PGTRPTAMGS FSSHMEFPR KRKGSDDSPS 60
QVEDGEHQVK MKAFREAHQ TEKRRRDKMN NLIELSAMI PQCNPMARKL DKLTVLRMAV 120
QHLRLSLKGLT NSYVGSNYRP SFLQDNELRH LILKTAEGFL PVVGCERGIK LFVSKSVSKI 180
LNYDQASLTG QSLFDPLHPK DVAKVKEQLS SFDISPREKL IDAKTGLQVH SNLHAGTRTV 240
YSSGRRSPFC RIKSKRISVK EHGCLFNSK KKEERKFYTI ECTGYLRSWP PNIVGMSEER 300
NSKDNSTFT CLVAIGRLQP YIVPQNSGEI NVKPTFTFTR FAVNGKVFVY DQRATALLGY 360
LPQELGTSFC YEYFHQDDHN NLTDKHKAVL QSKEKILIDS YKFRAGDGSF VTLKQWFSF 420
TNPWTKELEY IVSVNTLVLG HSEPGASFL PCSQSSEES SRQSCMSVPG MSTGTVLGAG 480
SIGTDIANEI LDQLRLQSSS YLDDSSPTGL MKDTHTVNCR SMSNKLFPF SPSEMGELEA 540
TRONQSTVAV HSHEPLSDG AOLDFDALCD NDDTAMAFM NYLEABGGIG DPGDFSDIQW 600
TL

Seq ID NO: 167 DNA sequence
Nucleic Acid Accession #: NM_014400
Coding sequence: 86-1126

60
65
70
75
80
85

1 11 21 31 41 51
GGTTACTCAT CCTGGGCTCA GGTAAGAGGG CCCGAGCTCG GAGGCGGCAC ACCCAGGGGG 60
GAGCCCAAGG GAGCAGGACG GAGCCATGGA CCCGCCCAGG AAGCAGGTG CCCAGGCCAT 120
GATCTGGACT GCAGGCTGGC TGCTGCTGCT GCTGCTTGC GGAGGAGCGC AGGCCCTGGA 180
GTGCTACAGC TGCGTGACGA AAGCAGATGA CGGATGCTCC CCGAACAGA TGAAGACAGT 240
GAAGTGGCGG CCGGGCGTGG ACGTCTGCAC CGAGGCGGTG GGGCGGTGG AGACCATCCA 300
CGGACAATTC TCGCTGGCAG TGCSGGGTTG CGGTTCGGGA CTCCCCGGCA AGAATGACCG 360
CGGCCCTGAT CTTCACGGGC TTCTGGCGTT CATCCAGCTG CAGCAATGCG CTCAGGATCG 420
CTGCAACGCC AAGCTCAACC TCACCTCGCG GCGGCTCGAC CGGCGAGGTA ATGAGAGTGC 480
ATACCGCGCC AACCGCGTGG AGTGCTACAG CTGTGTGGGC CTGAGCGGGG AGGCGTGCCA 540
GGGTACATCG CCGCCGGTGG TGAGCTGCTA CAACGCCAGC GATCATGTCT ACAAGGGCTG 600
CTTGGACGGC AAGCTCACCT TGACGCGCAG TAATGTGACT GTGTCTTTCG CTGTCCGGGG 660
CTGTGTCCAG GATGAATTTCT GCACTCGGGA TGGAGTAACA GGCCAGGGT TCACGCTCAG 720
TGGCTCCTGT TGCCAGGGGT CCGCTGTGTA CTCTGACCTC CGCAACAAGA CCACTTCTC 780
CCCTCGAATC CCACCCCTTG TCCGGCTGCC CCCTCCAGG CCAACGAGT TGGCCTCAAC 840
CACATCTGTC ACCACTTCTA CCTCGGCCCC AGTGAGACCC ACATCCACCA CCAACCCAT 900
GCCAGGCCCA ACCAGTCAGA CTCGAGACA GGGAGTAGAA CAGAGGCCCT CCGGGATGA 960
GGAGCCCGAG TTGACTGGAG GCGCGCTGG CCACAGGAC CGCAGCAATT CAGGGCAGTA 1020
TCCTGCAAAA GGGGGGCCCC AGCAGCCCA TAATAAAGGC TGTGTGGCTC CCACAGCTGG 1080
ATTGGCAGCC CTCTCTGTGG CCGTGGCTGC TGGTGTCTTA CTGTGAGCTT CTCACCTGG 1140
AAATTTCCCT CTCACCTACT TCTCTGGCCC TGGGTACCCC TCTTCTATC ACTTCTCTGT 1200
CCCAACACTG GACTGGGCTG GCCAGCCCC TGTTTTCCA ACATTCCCCA GTATCCCCAG 1260
CTTCTGCTGC GCTGTTTGGC GGCTTTGGGA AATAAATAT CGTGTATAT ATTCTGGCAG 1320
GGGTGTTCTA GCTTTTGGAG GACAGCTCCT GTATCTTCT CATCCTGTG TCTCGGCTG 1380
TCCTCTTGTG ATGTTAGGAC AGAGTGAGAG AAGTCAGCTG TCACGGGGAA GGTGAGAGAG 1440
AGGATGCTAA GCTTCTACTT CACTTTCTCC TAGCCAGCCT GGACTTTGGA GCGTGGGGTG 1500
GGTGGGACAA TGCTCTCCCA CTCTAAGCAC TGCTCCCTCT ACTCCCGCA TCTTTGGGGA 1560
ATCGGTTCCT CATATGTCTT CCTTACTAGA CTGTGAGCTC CTGAGGGGCA GGGACCGTGC 1620
CTTATGCTCG TGTGTGATCA GTTCTGGCA CATAAATGCC TCAATAAGA TTTAATTACT 1680

TTGTATAGTG AAAAAAA

Seq ID NO: 168 Protein sequence:
Protein Accession #: NP_055215

1	11	21	31	41	51	
MDPARKAGAQ	AMINTAGWLL	LLLLRGGQAQ	LECYSCVQKA	DDGCSPNKMK	TVKCAPGV DV	60
CTEAVGAVET	IHGQFSLAVX	GCGSGLPGKN	DRGLDLHGLL	AFIQLQQCAQ	DRCNAKINLT	120
SRALDPAGNE	SAYPPNGVEC	YSCVGLSREA	CQGTSPFVVS	CYNASDRHYK	GCFDGNVLT	180
AANVTVSLFV	RGCVDDEFCT	RDGVTGPGFT	LSGSCCQGSR	CNSDLRNKTY	FSPRIPLVR	240
LPPPEPTTVA	STTSVTSTTS	APVPTSTTK	PMPAPTSQTP	RQGVHEASR	DEEPLRTGGA	300
AGHQDRSNSG	QYPAKGPGQQ	PENKGCVAPT	AGLAALLAV	AAGVLL		

Seq ID NO: 169 DNA sequence
Nucleic Acid Accession #: NM_006875
Coding sequence: 186-1190

1	11	21	31	41	51	
GAATTCGGCA	CGAGCGCGCG	GCGAATCTCA	ACGCTGCGCC	GTCTCGGGCC	GCTTCGGGGC	60
CACCAAGTTC	TCTGCTTTCC	ACCCTGGCGC	CCCCCAGCCC	TGGCTCCCCA	GCTGCGCTGC	120
CCCGGGCGTC	CACGCCCTGC	GGGCTTAGCG	GGTTCAGTGG	GCTCAATCTG	CGCAGCGCCA	180
CTCCCATGTT	GACCAAGCTC	CTACAGGGGC	CTCCGCGCCC	CCCGGGGACC	CCCACGCCGC	240
CGCCAGGAGG	CAAGGATCGG	GAAGCGTTCC	AGGCGAGSTA	TGCACTCGGC	CCCCTCTGG	300
GTAAGGGGGG	CTTTGGCACC	GTCTTCGCAG	GACACCCCTT	CACAGATCGA	CTCCAGGTGG	360
CCATCAAAGT	GATTCCCGGG	AATCGTGTGC	TGGGCTGCTC	CCCCTTGTCA	GACTCAGTCA	420
CATGCCCACT	GGAAGTCGCA	CTGCTATGGA	AAGTGGGTGC	AGGTGGTGGG	CACCTGGCGG	480
TGATCCGCTC	GCTTGACTGG	TTTGAGACAC	AGGAAGGCTT	CATGCTGGTC	CTCGAGCGGC	540
CTTTGCCCGC	CCAGGATCTC	TTTGAATATA	TCAAGAGAAA	GGGCCCACTG	GGTGAAGGCC	600
CAAGCCGCTG	CTTCTTTGGC	CAAGTAGTGG	CAGCCATCCA	GCACTGCCAT	TCCCGTGGAG	660
TTGTCCATCG	TGACATCAAG	GATGAGAAC	TCCTGATAGA	CCTACGCCGT	GGCTGTGCCA	720
AACCTCATGA	TTTGGTTTCT	GGTGCCCTGC	TTTATGATGA	ACCCTACACT	GACTTTGATG	780
GGACAAGGGT	GTACAGCCCC	CCAGAGTGG	TCTCTCGACA	CCAGTACCAT	GCACTCCCGG	840
CCACTGTCTG	GTACTGGGGC	ATCTCTCTCT	ATGACATGCT	GTGTGGGGAC	ATTCCCTTTG	900
AGAGGGACCA	GGAGATTCTG	GAAGCTGAGC	TCCACTTCCC	AGCCCATGTC	TCCCAGACT	960
GCTGTGCGCT	AATCCGCGCG	TGCTTGGCCC	CCAAACCTTC	TTCCCGACCC	TCACTGGAAG	1020
AGATCTCTGT	GGACCCCTGG	ATGCAAAAC	CAGCCGAGGA	TGTTACCCCT	CAACCCCTCC	1080
AAAGGAGGGC	CTGCCCTTTT	GGCCTGTGTC	TTGCTACCCCT	AAGCCTGGCC	TGGCCTGGCC	1140
TGGCCCCCAA	TGGTCAGAA	AGCCATCCCA	TGGCCATGTC	ACAGGGATAG	ATGGACATTT	1200
GTTGACTTGG	TTTTACAGGT	CATTACCACT	CATTAAAGTC	CAGTATTACT	AAGTAAGGG	1260
ATTGAGGATC	AGGGGTAGAA	AGACATAAAC	CAAGTTTGCC	CAGTTCCCTT	CCCAATCCTA	1320
CAAAAGGAGC	TTCTCCCGAG	AACCTGTGGT	CCCTGATTTT	GGAGGGGGAA	CTTCTTGCTT	1380
CTCATTTTGC	TAAAGAAAGT	TATTTTGGTG	AAGTTGTGTC	CATTTTGAGC	CCCGGAGCTC	1440
TTATTTTGAT	GATGTGTGAC	CCCACATGCG	CACCTCCTAC	TACCACCACA	CAAACCTAGT	1500
TCATATGCTT	TTACTTGGGC	AAGGGTGCTT	TCCTTCCAAT	ACCCAGTAG	CTTTATTTTT	1560
AGTAAAGGGA	CCCTTTCCCG	TAGCCTAGGG	TCCCATATTG	GGTCAAGCTG	CTTACCTGCC	1620
TCAGCCCGAG	ATTTTATTAT	TTGGGGGAGG	TAACTCCCTG	TTGTTACCCC	AAGGCTTCTT	1680
TTTTTTTATT	TTTTTTTATT	GGTGAGGGGA	CCCTACTTTG	TTATCCCAAG	TGCTCTTATT	1740
CTGTGAGAAA	GAGCTTAATT	TCCATAATTT	GGGAAGGAAT	GGAAGATGGA	CACCACCGGA	1800
CACCACCAGA	CAATAGGATG	GGATGGATGG	TTTTTTGGGG	GATGGGCTAG	GGGAAATAAG	1860
GCTTGTCTGT	TGTTTTCTGT	GGGCGCTCCC	TCCAATTTTG	CAGATTTTGG	CAACCTCCTC	1920
CTGAGCCGGG	ATTGTCCCAAT	TACTAAATAT	TAAATAATCA	CGTATTGTGG	GGAGGGGAGT	1980
TCCAAGTGTG	CCCTCCTTTT	TTTTCTCTGC	TGGATTATTT	AAAAAGCCAT	GTGTGGAAC	2040
CCACTATTTA	ATAAAAGTAA	TAGAATCAGA	AAAAAATAAA	AAAAAATAAA		

Seq ID NO: 170 Protein sequence:
Protein Accession #: NP_006866

1	11	21	31	41	51	
MLTKPLQGGP	APPGTPTPPP	GGKDREAFEA	EYRLGPLLGG	GGFGTVFAGH	RLTDRLQVAI	60
KVIPRNRLVG	WSPLSDSVTC	PLEVALLWKV	GAGGGHPGVI	RLLDWFETQE	GFMLVLERPL	120
PAQDLFDYIT	EKGPLGEGPS	RCFPQGVVAA	IQHCHSRGVV	HRDIKDENIL	IDLRRGCARK	180
IDFGSGALLH	DEPYTFDDGT	RVYSPPPEWIS	RHQYHALPAT	VNSLIGILLVD	MVCGDIPFER	240
DQRILEARLH	PPAHVSPDCC	ALIRRCCLAPK	PSSRPSLEBI	LLDPWMQTPA	EDVTPQPLQR	300
RPCPFFGLVLA	TLSLAWPGLA	PNGQKSHFMA	MSQG			

Seq ID NO: 171 DNA sequence
Nucleic Acid Accession #: NM_003646
Coding sequence: 89..2875

1	11	21	31	41	51	
GCGGCGGGGA	GCGGGCGTGC	TGAGCCCCCG	CGGCGGGCCC	GGCATGGGCG	TCTCCCGCGG	60
GCCTTCGCGC	GGCGGGGGCT	AGGGCCGGAT	GGAGCGCGCG	GACGGTAGCC	CCGAGGCGCG	120
GAGCAGCGAC	TCCGAGTCGG	CTTCCGCTCC	GTCCAGCGCG	TCCGAGCGCG	ACGCGGTGCC	180
CGAGCGGGAC	AAGGCGCCCG	GGCGACTCAA	CAAGCGGCGC	TTCCGGGGGC	TGCGGCTCTT	240
CGGGCAGCAG	AAAGCCATCA	CCAAGTCGGG	CCTCCAGCAC	CTGGCCCCCC	CTCCGCGCAC	300
CCCTGGGGCC	CGTGAGCAGG	AGTCAGAGCG	GCAGATCCGG	AGTACAGTGG	ACTGGAGCGA	360
GTACAGCGCA	TATGGGAGCG	ACATCTGGTT	CGAGACCAAC	GTGTCCGGGG	ACTTCTGCTA	420
CGTTGGGGAG	CAGTACTGTG	TAGCCAGGAT	GCTGAAGTCA	GTGTCTCGAA	GAAAGTGCGC	480
AGCCTGCAAG	ATTGTGGTGC	ACACGCCCTG	CATCGAGCAG	CTGGAGAAGA	TAAATTTCCG	540
CTGTAAGCCG	TCCTTCCGTG	AATCAGGCTC	CAGGAATGTC	CGCGAGCCAA	CCTTTGTACG	600
GCACCACTGG	GTACACAGAC	GACGCCAGGA	CGGCAAGTGT	CGGCACTGTG	GGAGGAGATT	660
CCAGCAGAAG	TTCACCTTCC	ACAGCAAGGA	GATTGTGGCC	ATCAGCTGCT	CGTGGTGCAA	720

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

GCAGGCATAC CACAGCAAGG TGTCTGCTT CATGCTGCAG CAGATCGAGG AGCCGTGCTC 780
GCTGGGGGTC CACGCGAGCG TGGTCATCC GCCCACTGG ATCCTCGCG CCGGAGGCC 840
CCAGAATACT CTGAAGCAA GCAAGAAGAA GAAGAGGGCA TCCTTCAAGA GGAAGTCCAG 900
CAAGAAAGGG CCTGAGGAGG GCGCTGGAG ACCCTTCATC ATCAGGCCCCA CCCCCTCCCC 960
GCTCATGAG CCTCTGCTGT TGTITGTGAA CCCCAGAGT GGGGGCAACC AGGGTGCAAA 1020
GATCATCCAG TCTTCTCTCT GGTATCTCAA TCCCAGCAA GTCTTCGACC TGAGCCAGGG 1080
AGGGCCCAAG GAGGCGCTGG AGATGTACCG CAAAGTGCAC AACCTGCGGA TCCTGGGGTG 1140
CGGGGGGAC GGCACGGTGG GCTGGATCCT CTCACCCCTG GACCAGCTAC GCCTGAAGCC 1200
GCCACCCCTT GTTCCATCC TCGCCCTGGG TACTGGCAAC GACTTGGCCC GAACCTCAA 1260
CTGGGGTGGG GGCTACACAG ATGAGCCTGT GTCCAAGATC CTCTCCACAG TGGAGGAGGG 1320
GAACTGTGTA CAGCTGGACC GCTGGGACCT CCACGCTGAG CCCAACCCCG AGGCAGGGCC 1380
TGAGGACCGA GATGAAGGCG CCACGACCG GTTGCCCTG GATGTCTTCA ACAACTACTT 1440
CAGCTGGGCG TTTGAGCGCC ACCTCACCTT GGAGTTCAC GAGTCTCGAG AGGCCAACCC 1500
AGAGAAATTC AACAGCGCTT TCGGAATAA GATGTTCTAC GCGGGGACAG CTTTCTCTGA 1560
CTTCTGATG GGCAGCTCCA AGGACCTGGC CAAGCACATC CGAGTGGTGT GTGATGGAAT 1620
GGACTTGACT CCCAAGATCC AGGACCTGAA ACCCCAGTGT GTTGTCTTCC TGAATCATCC 1680
CAGTACTGT GCGGGCAGCA TCCCTGGGG CCACCTGGG GAGCACCCAG ACTTTGAGCC 1740
CCAGCGGCAT GACGAGCGGT ACCTCGAGGT CATTGGCTTC ACCATGACGT CGTGGCCGCG 1800
GCTGCAGGTG GCGGACACAG GCGAGCGGCT GACGAGTGT GCGAGGTGT TGCTCACCAC 1860
ATCCAAGGCG ATCCCGGTGG AGGTGGATGG CGAGCCCTGC AAGCTTGAG CCTCAGCAT 1920
CCGATCGCC CTGCGCAACC AGGCCACCAT GGTGCAAGG GCCAAGCGGC GGAGCGCGCG 1980
CCCCCTGCAC AGCAGCCAGC AGCCGCTGCC AGAGCAGTTG GGCATCCAG TGAGTCGCGT 2040
CAGCATGCAC GACTATGAGG CCCTGCACTA CGACAAGGAG CAGCTCAAG AGGCCTCTGT 2100
GCGCTGGGG ACTGTGGTGG TCCAGGAGA CAGTGACCTA GAGCTCTGCC GTGCCACAT 2160
TGAGAGACTC CAGCAGGAGC CCGATGGTGG TGGAGCCAAG TCCCAGCAT GCCAGAACT 2220
GTCCCCAAG TGGTGTCTCC TGGACGCCAC CACTGCCAGC CGCTTCTACA GGATCGACCG 2280
AGCCAGGAG CACCTCAACT ATGTGACTGA GATGSCACAG GATGAGATT ATATCCTGGA 2340
CCCTGAGCTG CTGGGGGCAT CCGCCCGGCC TGACCTCCA ACCCCACTT CCCCCTCTCC 2400
CACTCACCC TGCTCAACCA CCGCCCGGTC ACTGCAAGGG GATGCTGCAC CCCCCTCAAG 2460
TGAAGAGCTG ATTGAGGAG CCAAGAGGAA CGACTTCTGT AAGCTCCAG AGCTGCAACG 2520
AGCTGGGGG GACCTCATGC ACCGAGACGA CGAGAGTGGC AGCTCTCTGC ACCACGAGT 2580
CAGCACTGGC AGCAAGGATG TGGTCCGCTA CCTGCTGAG CACGCCCCCG CAGAGATCCT 2640
TGATGGGGT GAGGAAACG GGGAGACCTG TTTGCACCAA GCAGCGGCC TGGGCCAGCG 2700
CACCATCTCG CACTACATCG TGGAGGCGCG GGCCTCGCTC ATGAAGACAG ACCAGCAGGG 2760
CGACACTCCC CGGCGCGGCG CTGAGAGGC TCAGGACACC GAGCTGGCGC CCTACCTGGA 2820
GAACCGCAG CACTACCAGA TGATCCAGCG GGAGGACCAG GAGACGGCTG TGTAGCGGGC 2880

Seq ID NO: 172 Protein sequence:
Protein Accession #: NP_003637

1 11 21 31 41 51
MEPRDGSPEA RSSDSESASA SSSGSRDAG PEPDKAPRRL NKRRFPGLRL FGHRKAITKS 60
GLQHLAPPPP TPAGPCESE RQIRSTVDWS ESATYGEHW PETNVSGDFC YVGEQYCVAR 120
MLKSVSRKRC AACKIVVHTP CIEGLEKINF RCKPSFRESG SRNVREPTFV RHHVHRRRQ 180
DGKCRHCGKG PQQKPTFHSK EIVAISSWC KQAYHSHVSK FMLQJIEPC SLGVAHAVVI 240
PPTWILRARR PONTLKASK KKRASFKRKS SKKGPEGRW RPFIIPTPS PLMKPLLVFV 300
NPKSGGNQGA KIIQSFLWYL NPKQVFDLSQ GGPKEALEMY RKVHNLRLA CGGDGTVGWI 360
LSTLDQLRLK PPPFVAILPL GTGNDLARTL NWGGGYTDEP VSKILSHVEE GNVVQLDRWD 420
LHAEPNPEAG PEDRDEGATD RLPLDVFNRY FSLGFDHVT LEPHESREAN PEKFNRRFRN 480
KMFYAGTAFS DFLMSSKDL AKHIRVVDG MDLTPKIQDL KPQCVFLNI PRYCACTMPW 540
CHPGEHDFE PQRHDDGYLE VIGFTMTSLA ALQVGGHER LTQCREVLT TSKAIPQVD 600
GSPCKLAASR IRIALRNQAT MVQAKRRA APLHSDQPV PEQLRIQVSR VSMHDEALH 660
YDKEQLKEAS VPLGTVVVPG DSDLELCRAH IERLQEPDG AGAKSPTQK LSPKWCFLDA 720
TTASRFYRID RAQEHNLVYT EIAQDEIYIL DPELLGASAR PDLPTPTSP PLSPSPPTPR 780
SLQGDAAFPQ GEELIEAAKR NDFCKLQELH RAGGDLMRD EQSRTLLHHA VSTGSKDVR 840
YLLDHAPPEI LDAVEENGET CLHQAAALGQ RTICHYIVEA GASLMKTDQ GDTFRQRAEK 900
AQDTELAAYL ENRQHYQMIG REDQETAV

Seq ID NO: 173 DNA sequence
Nucleic Acid Accession #: AF232772
Coding sequence: 1-1662

1 11 21 31 41 51
ATGCCGGTGC AGCTGACGAC AGCCCTGCGT GTGGTGGGCA CCAGCCTGTT TGCCCTGGCA 60
GTGCTGGGTC GCATCGCTGC AGCCTATGTG ACGGGCTACC AGTTCATCCA CACGGAAG 120
CACTACCTGT CCTTCGCTGT GTACGGCGCC ATCCTGGGCC TGCACTGCTC CATTGAGAC 180
CTTTTGTGCT TCCTGGAGCA CCGGCGCATG CGAGCTGCCG GCCAGGCCCT GAAGTGGCCC 240
TCCCGCGGCG GGGGCTCGGT GGCACTGTGC ATTGCGCAT ACCAGGAGGA CCTGACTAC 300
TTGGCGAAGT GCCTGCGCTC GGCCAGCGC ATCTCCTCC CTGACCTCAA GGTGGTCATG 360
GTGGTGCATG GCAACGSCCA GGAGGACGCC TACATGCTGG ACATCTTCCA CGAGGTGCTG 420
GGCGGCACCG AGCAGGCGCG CTCTTTGTG TGGCGCAGCA ACTTCCATGA GGCAGGCGAG 480
GGTGAGACGG AGGCAGCCTC GCAGGAGGGC ATGGACCGTG TGGGGGATGT GGTGCGGGCC 540
AGCACCTTCT CGTGCATCAT CGAGAAGTGG GGAGGCAAGC CGAGGTGATG GTACAAGGCC 600
TTCAAGGCCCT TCGGCGATTC GGTGGACTAC ATCCAGGTGT GCGACTCTGA CACTGTGCTG 660
GATCCAGCCT GCACCATCGA GATGCTTCGA GTCTGGAGG AGGATCCCCA AGTAGGGGGA 720
GTGGGGGAG ATGTCCAGAT CCTCAACAG TACGACTCAT GGATTTCTCT CCTGAGCAGC 780
GTGCGGTACT GGTATGCTCT CAACGTGGAG CGGGCTGCCC AGTCTACTT TGGCTGTGTG 840
CAGTGTATTA GTGGGCGCTT GGGCATGTAC CGCAACAGCC TCCTCCAGCA GTTCTGGAG 900
GACTGTATCC ATCAGAAGTT CCTAGGCAGC AAGTGCAGCT TCGGGGATGA CCGGACCTC 960
ACCAACGAG TCCTGAGCCT TGGCTACCGA ACTAAGTATA CCGCGGCTC CAAGTGCCTC 1020
ACAGAGACCC CCACTAAGTA CCTCGGTTGG CTCAACAGC AAACCGCTG GAGCAAGTCT 1080
TACTTCGCGG AGTGGCTCTA CAACTCTCTG TGGTTCATA AGCACCCTC CTGGATGACC 1140
TAGAGTCAG TGGTCAAGG TTTCTTCCCC TTCTTCTCA TTGCCAGGT TATACAGCTT 1200
TTCTACCGG GCGGCATCTG GAACATCTC CTCTCTCTG TGACGGTGA GCTGGTGGGC 1260
ATTATCAAG CCACCTAGCG CTGCTTCTCT CCGGGCAATG CAGAGATGAT CTTCTATGTC 1320

5	CTCTACTCCC	TCTCTATAT	GTCCAGCCTT	CTGCCGGCCA	AGATCTTTGC	CATTGCTACC	1380
	ATCAACAAAT	CTGGCTGGGG	CACCTCTGGC	CGAAAAACCA	TTGTGGTGAA	CTTCATTGGC	1440
	CTCATTCTCG	TGTCCATCTG	GGTGGCAGTT	CTCCTGGAGG	GGCTGGCCTA	CACAGCTTAT	1500
	TGCCAGGACC	TGTTCACTGA	GACAGAGCTA	GCCTTCTCTG	TCTCTGGGGC	TATACCTGTAT	1560
	GGCTGCTACT	GGGTGGCCCT	CCTCATGCTA	TATCTGGCCA	TCATCGCCCG	GGGATGTGGG	1620
	AAGAAGCCGG	AGCAGTAGAG	CTTGGCTTTT	GCTGAGGTGT	GACATGGGCC	CCAAGCAGAG	1680
	CGGTAAGAT	CAATGGGTA	AGGGAGGAA	GGGAATGGA	AGAGAAAGA	CAGGGTGGGA	1740
	GGGAGGAGG	AGTGTCTGT	TTTAGTCTCT	TAATGTGCA	AAGGACAAAT	CTAAATGCA	1800
10	AAGAACGCTG	ATGTAGTAG	GCCTGACAGC	TCTGTTTAGA	GGAGGCAACA	CTGATCCCCC	1860
	AGATGCAGAG	CTGACGGGGA	TTCTGTGTTT	TCAGACTGCC	TGTCTGCTTG	CATCTGCACA	1920
	TAGGCAGTAG	CCTCCTCCTG	GGCTCCAGAG	GGCACTCAGA	AGTTGTGCTA	AAACCAAGTTA	1980
	AGTCCCATTC	AGTGGCAACT	TGTATAGGT	ACCTGAGTGA	CGGCAACCTG	CGGAAGGAGG	2040
	TTCTCCAGC	CCATCTGAAC	ACAACCAGAG	GTGGCAGGAG	AATTTCTACT	GAGCGAGGTG	2100
	GGCGGTTAG	TGTATGTAC	CCCCACCCCA	CCCATAACTA	GTCTCAATG	CAATAAGATT	2160
15	GGCGGTGAGA	TACAAGGCC	AGAAGCCTGA	TCTTTGGGCA	TCAGAAAACA	GGGTCCAGGA	2220
	ATGGTCTTTT	ATGTGAGATA	CCCCACTCCA	CATCAACATT	CCAGGATGA	GCCAAACCAG	2280
	CAGGGAGTTA	GCATGAAC	GCTTTTAAAA	GTGCACATTA	AAAAGGAAAG	TTTGCCAGGA	2340
	GGAAACAAAGA	GATTGTGTG	GTGCTAAAGG	AGGCCATAAG	CTACACAGAG	GCCTTGGGTG	2400
20	TTCCACCTGG	AAACTGCTCA	GACGTCTAGA	TGGGTTCTTA	GCTTGTCTGT	GATCTCTGCT	2460
	GGGAGATAAA	AAAGATTAA	CCCCAACATG	TCAGAAAAG	AAGTGAAGCT	TTGGGTATT	2520
	TAACCTGTAT	ACTCTTGAAT	TCCTCTCAAA	TTCACTCTG	ATCTGAGGCT	AAGACACACT	2580
	CCCCACTTCA	CTTTCTTCAA	AGCCACATTT	TTTGAGGTAT	CACTGCAGTC	ACCTCTTCTA	2640
	CCCTCATCAT	CATAGTAAAG	GTTTTCAAGG	TGGCAATTGG	GGCGGAGCCC	CGGCTTCTTA	2700
25	TAGAAGCTTC	AGCAGGAGGC	AAGCGTGTTC	TCAGCACATA	TGGGAAGTAT	GAGGAGCCTC	2760
	TGATCAAAAT	TACAGGAATC	TTGGAGCTGC	TTGGACGGAT	TCCTTGGCAG	CGGGTTAGC	2820
	ATGTGTGACT	TTCAAGCTAC	TGTTCTTGAC	AATCATCTCC	AATGGAAAGC	TTTTCAGTGT	2880
	TCCCAAGGTG	AACTCTCAA	TCCAAATG	TTATCTTGA	GACCATCCAT	TCTCTCAGT	2940
	GGCTTCTCCA	GGGAATCTT	ACAGCCAAGT	TGTGACAGTC	ACTGCATTG	CCTGCTTCTT	3000
30	TCAGAAACCC	AAACTAGGAG	ATGAACTGG	TTCTACATC	CTAAGGTTCT	TGCTTTCTCT	3060
	CTCATGCCCT	TGGTGGCTGT	TTTGGCTGT	TTTCCCTCTG	CTGCTTTTGG	GGAAATGAGG	3120
	GAAGCCATTT	TCCAAGTGAC	TTGCAATCCA	GGCTGTCTC	AGCGTTTGA	GTTTAAAACC	3180
	TGGATCCTG	ACTAAGCCTT	TGACTTAAGG	GTTGCTTGT	TGCCCTCCAA	ATGCTCTTTC	3240
	TCAAAGGGGC	CAACTAAACC	GTGCAGAAC	AGCACTAAGG	TGGACAGCAG	ACAAGAGGGC	3300
35	AAGCCTCTAA	TGTACCAAGT	GCTTCTTACA	AAGACGCAAG	GTGTGCTCCG	AACCAAGAGT	3360
	GGGCAAAACC	TGGTGGCTTC	CTTCACTTCC	CAGCAACTCA	AGGGTTTTC	AAGTGTAGCT	3420
	AACAGTTGCC	ACATCACACA	GACCTCCAGT	TTCTGGTAAG	ACTGCTGTTT	GACATCAGAC	3480
	CCAACCCATT	GAAGGCTGGA	AGGCAGCAGG	CATTGTCTAA	GGCAGCTGAT	CCAGGCAATC	3540
	GTTCGTCTGG	CCAAGAAATT	AAACTATTTT	GAGCATTAGA	ATGGAGGAAA	TCCGTCAGC	3600
40	CAAGTGACAGA	GTTCAAGCTT	CGCTAAGGGC	TTGTTTTTCT	TCAGCATTTA	CTTGAAGATT	3660
	AATGTAGGAT	GACAGGCTCT	CCTGGCTGTC	CTACCATCAG	CTCTGCCTTG	CACTGTGCTC	3720
	GTCAACTTTC	CTCAATACAA	AAACAGGCAG	GTACAGGTAG	TGGGCTCACA	ACGTTTGACC	3780
	TOGACTGGTT	TTTCTAAGTT	ATTTTGTACA	TTTTTCAGCA	GCAAAACCAA	ACTGGGTCTT	3840
	CAGCTTTATC	CCCGTTTCTT	GCAAGGGAAG	AGCCTTTATA	CAATTGGACG	CATTTTGGTT	3900
45	TTTCTCTATT	GAGAATTCAA	ATCCTCTTTT	GTATTGTTTC	TACAATAATT	TGTAACATA	3960
	TTTATTTTAA	CCTGCTTTT	TTTTTTTTTT	TAATTTTCAG	GTCAAGTTT	TTATACTGCA	4020
	CTTATTGTGC	AAAAATAAGA	TTCTCACAT				

Seq ID NO: 174 Protein sequence:
Protein Accession #: AAF36984

50	1	11	21	31	41	51	
	MPVQLTTLALR	VVGTSFLFALA	VLGGILAAAYV	TGYQPIHTEK	HYLSFGLYGA	ILGLHLLIQS	60
55	LFAPLEHRRM	RRAGQALKLP	SPRRGSVALC	IAAYQEDPDY	LRKCLRSAQR	ISFPDLKVVV	120
	VVDGNRQEDA	YMLDIFHEVL	GGTEQAGFFV	WRSNFHEAGE	GETEASLQEG	MDRVDRVVRA	180
	STFSCIMQKW	GGKREVMYTA	FKALGDSVDY	IQVCDSDTVL	DPACTIEMLR	VLEEDPQVGG	240
	VGGDVQILNK	YDSWISFLSS	VRWMAFNVE	RACQSYFCGV	QCISGLPMY	RNSLLQQFLE	300
	DNYHQKPLGS	KCSFGDDRHL	TNRVLSLGR	TKYTARSKCL	TETPTKYLWR	LNQQTNRWSK	360
60	YFREMLYNSL	WFHKKHLMWT	YESVVTGFPF	FFLIATVIQL	FYRGRINWIL	LFLLTQVLQG	420
	IKATYACFL	RGNAMIFMS	LYSLLYMSSL	LPAKIFAIAT	INKSGWGTSG	RKTIVNFTIG	480
	LIPVSIWVAV	LLEGLAYTAY	QDLFSETEL	APLVSGAILY	GCYWVALLML	YLAI IARRCG	540
	KKPEQYSLAF	AEV					

Seq ID NO: 175 DNA sequence
Nucleic Acid Accession #: NM_000691
Coding sequence: 43..1404

70	1	11	21	31	41	51	
	CCAGGAGCCC	CAGTTACCGG	GAGAGGCTGT	GTCAAAGGCG	CCATGAGCAA	GATCAGCGAG	60
	GCCGTGAAGC	GCGCCCGCGC	CGCCTTCAGC	TCGGGCGAGG	CCCGTCCGCT	GCAGTTCCGA	120
	TTCCAGCAGC	TGGAGGCGCT	GCAGCGCCTG	ATCCAGGAGC	AGGAGCAGGA	GCTGGTGGGC	180
75	GGCTGCGCCG	CAGACCTGCA	CAAGAATGAA	TGGAACGCCT	ACTATGAGGA	GGTGGTGTAC	240
	GTCTAGAGG	AGATCGAGTA	CATGATCCAG	AAGCTCCCTG	AGTGGGCGCG	GGATGAGCCC	300
	GTGGAGAAGA	CGCCCCAGAC	TCAGCAGGAC	GAGCTCTACA	TCCACTCGGA	GCCACTGGGC	360
	GTGGTCTCTG	TCATTGGCAC	CTGGAACCTAC	CCCTTCAACC	TCACCATCCA	GCCCATGGTG	420
	GGCGCCATCG	CTGCAGGGA	CGCAGTGTC	CTCAAGCCCT	CGGAGCTGAG	TGAGAACATG	480
80	GCGAGCCTGC	TGGCTACCAT	CATCCCCCAG	TACCTGGACA	AGGATCTGTA	CCCAGTAATC	540
	AATGGGGGTG	TCCCTAGAGC	CACGGAGCTG	CTCAAGGAGA	GGTTCGACCA	TATCCTGTAC	600
	ACGGGCAGCA	CGCGGGTGGG	GAAGATCATC	ATGACGGCTG	CTGCCAAGCA	CCTGACCCCT	660
	GTACGCTGGG	AGCTGGGAGG	GAAGAGTCCC	TGCTACGTGG	ACAAGAAGCT	TGACCTGGAC	720
	GTGGCCTGGC	GACGATCGCG	CTGGGGGAAA	TTATGAAACA	GTGGCCAGAC	CTGCTGGGCC	780
	CCAGACTACA	TCTCTGTGTA	CCCTCGATC	CAGAACCAAA	TTGTGGAGAA	GCTCAAGAAG	840
85	TCACTGAAAG	AGTTCTACGG	GGAAGATGCT	AAGAAATCCC	GGGACTATGG	AAGAATCATT	900
	AGTGCCCGGC	ACTTCCAGAG	GGTGTATGGC	CTGATTGAGG	GCCAGAAGGT	GGCTTATGGG	960
	GGCACCGGGG	ATGCGGCCAC	TGCTACATA	GCCCCACCA	TCCTACGGA	CGTGAGCCCC	1020

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

CAGTCCCCGG TGATGCAAGA GGAGATCTTC GGGCCTGTGC TGCCCATCGT GTGCGTGGGC 1080
AGCCTGAGAG AGGCCATCCA GTTTCATCAAC CAGCGTGAGA AGCCCTGGC CCTCTACATG 1140
TTCTCCAGCA ACGACAAGGT GATTAAGAAG ATGATTGCAG AGACATCCAG TGGTGGGGTG 1200
GCGGCCAAGC ATGTGTCATG CCACATCACC TTGCACTCTC TGCCCTTCGG GGGCGTGGGG 1260
AACAGCGGCA TGSGATCCTA CCATGGCAAG AAGAGCTTCG AGACTTTCTC TCACCGCGGC 1320
TCTTGCCCTG TGAGGCCTCT GATGAATGAT GAAGGCCTGA AGGTCAGATA CCCCCGAGC 1380
CCGGCCAAAG TGACCCAGCA CTGAGGAGGG GTTGCTCCGC CTGGCCTGGC CATACTGTGT 1440
CCCATCGGAG TGCGGACCAC CCTCACTGGC TCTCCTGGCC CTGGAGAATC GGTCTGCGAG 1500
CCCCAGCCCA GCCCAGCTCC TCTGCTGACC TGCTGACCTG TGACACCCCC ACTCCACAT 1560
GGGCCAGGC CTCACCATTC CAAGTCTCCA CCCCCTTTCTA GACCAATAAA GAGACAAATA 1620
CAATTTTCTA ACTCGG

Seq ID NO: 176 Protein sequence:
Protein Accession #: NP_000682

1 11 21 31 41 51
MSKISEAVKR ARAAFSSGRT RPLQFRFQQL EALQRLIQEQ EQELVGALAA DLHKNEWNAY 60
YEEVVVYLEE IEYMIQKLPE WAADEPVEKT PQTQDELYI HSEPLGVVLV IGTWNYPFNL 120
TIQPMVGAIA AGNAVVLKPS ELSENMASSL ATIIPOYLDK DLYPVINGGV PETTELLKER 180
FDHILYTGST GVGRIIMTAA AKHLTPVTLE LGGKSPCYVD KNCDDLVA CR RIAWCKFMNS 240
GQTCAVPDYI LCDPSIQNI VEKLLKSLKE FYGEDAKKSR DYGRISARH FORVMGLIEG 300
QKVAYGTGD AATRYIAPT I LTDVDPQSPV MQEEIFGFVL PIVCVRSLEE AIQPINQREK 360
PLALYMFSSN DKVIKMIAE TSSGGVAAND VIVHITLHSL PFGGVNSGM GSYHGKKSFE 420
TFSHRRSCLV RPLMNDEGLK VRYPPSPAKM TQH

Seq ID NO: 177 DNA sequence
Nucleic Acid Accession #: NM_001067.1
Coding sequence: 108-4703

1 11 21 31 41 51
CTAACCGAGC GCGCTCTGTG GAGAAGCGGC TTGGTGGGGG GTGGTCTCGT GGGTCTCTGC 60
CTGTTTAGTC GCTTTCAGGG TTCTTGAGCC CCTTCAGCAC CGTCACCATG GAAGTGTAC 120
CATTGACGCC TGTAATGAA AATATGCAAG TCACAAAAT AAGAAAAAT GAAGATGCTA 180
AGAAAAAGCT GTCTGTGAA AGAATCTATC AAAAGAAAAC ACAATTGGAA CATATTTTGC 240
TCCGCCAGCA CACCTACATT GGTCTGTGG AATTAGTGAC CCAGCAAATG TGGGTTTACG 300
ATGAAGATGT TGGCATTAA TATAGGGAAG TCACTTTGT TCTGTGTTTG TACAAAATCT 360
TTGATGAGAT TCTAGTTAAT GCTGCGGACA ACAACAAAG GGACCCAAAA ATGCTTTGTA 420
TTAGAGTCAC AATTGATCCG GAAAACAAT TAATTAGTAT ATGGAATAAT GGAAGAGTA 480
TTCTGTGTGT TGAACACAAA GTTGAAGAAG TGTATGTCCC AGCTCTCATA TTTGACAGC 540
TCCTAATCTC TAGTAACAT GATGATGATG AAAAGAAAGT GACAGGTGGT CGAAATGGCT 600
ATGGAGCCAA ATTGTGTAAC ATATTACGTA CCAAAATTAC TGTGGAACAA GCCAGTAGAG 660
AATACAAGAA AATGTTCAAA CAGACATGGA TGGATAATAT GGGAGAGAGT GGTGAGATGG 720
AACTCAAGCC CTTCAATGGA GAAGATTATA CATGTATCAC CTTTCAGCCT GATTGTCTA 780
AGTTTAAAT GCAAGCGCTG GACAAAGATA TTGTGCACT AATGGTCAGA AGAGCATATG 840
ATATTGCTGG ATCCACCAA GATGTCAAAG TCTTTCTTAA TGGAAATAAA CTGCCAGTAA 900
AAGGATTGCG TAGTTATGTG GACATGTATT TGAAGGACAA GTTGGATGAA ACTGTTAAT 960
CCTTGAAAGT AATACATGAA CAAGTAAACC ACAGSTGGGA AGTGTGTTTA ACTATGAGTG 1020
AAAAAGGCTT TCAGCAAAT AGCTTTGTCA ACAGCATTGC TACATCCAAG GGTGGCAGAC 1080
ATGTTGATTA TGTAGCTGAT CAGATTGTGA CTAACCTTGT TCGATTGTG AAGAAGAAGA 1140
ACAAGGGTGG TGTGCGATG AAAGCACATC AGGTGAAAAA TCACATGTGG ATTTTGTAA 1200
ATGCCCTAAT TGAAGACCCA ACCCTTGACT CTCAGACAAA AGAAAAACATG ACTTTACAAC 1260
CCAAGAGCTT TGGATCAACA TGCCAATTGA GTGAAAAATT TATCAAAGCT GCCATTGGCT 1320
GTGATATTGT AGAAGACATA CTAACCTGGG TGAAGTTTAA GGCCCAAGTC CAGTTAAACA 1380
AGAAAGTGTG AGCTGTAAAA CATAATAGAA TCAAGGGAAT TCCCAAACTC GATGATGCCA 1440
ATGATGCAAG GGGCCGAAAC TCCACTGAGT GTACGCTTAT CCTGACTGAG GGAGATTGAG 1500
CCAAAACCTT GGCTGTTTCA GGCCTTGGTG TGGTTGGGAG AGACAAATAT GGGGTTTTCC 1560
CTCTTAGAGG AAAAATACTC AATGTTGAG AAGCTTCTCA TAAGCAGATC ATGGAATAATG 1620
CTGAGATTAA CAATATCATC AAGATTGTGG GTCTTCAGTA CAAGAAAAAC TATGAAGATG 1680
AAGATTCAAT GAAGACGCTC CGTTATGGGA AGATAATGAT TATGACAGAT CAGGACCAAG 1740
ATGGTCCCA CATCAAGGCC TTGCTGATTA ATTTTATCCA TCACAACCTG CCTCTCTTC 1800
TGOGACATCT TTTCTGGAG GAATTTATCA CTCCCATGT AAAGGTATCT AAAAACCAAGC 1860
AAGAATGGC ATTTTACAGC CTCTCTGAAT TTGAAGAGTG GAAGAGTTCT ACTCCAAATC 1920
ATAAAAAATG GAAAGTCAAA TATTACAAAG GTTGGGCAC CAGCACATCA AAGGAAGCTA 1980
AAGAATACCT TGCAGATATG AAAAGACATC GTATCCAGTT CAAATATTCT GGTCTCTGAG 2040
ATGATGCTGC TATCAGCCTG GCCTTTAGCA AAAACAGAT AGATGATCGA AAGGAATGGT 2100
TAACATAATT CATGGAGGAT AGAAGACAAC GAAAGTTACT TGGGCTTCTT GAGGATTACT 2160
TGATGGACA AACTACCACA TATCTGACAT ATAATGACT CATCAACAAG GAACTTATCT 2220
TGTTCTCAA TCTTGATAAC GAGAGATCTA TCCCTTCTAT GGTGGATGGT TTGAAACAG 2280
GTGAGAGAAA GGTTTTGTGT ACTTGCTTCA AACGGAATGA CAAGCGAGAA GTAAAGGTTG 2340
CCCCAATTAG TGGATCAGTG GCTGAAATGT CTCTTATCA TCATGGTGAG ATGTCATAA 2400
TGATGACCAT TATCAATTG GCTCAGAAAT TTGTGGGTAG CAATAATCTA AACCTCTTGC 2460
AGCCCATTTG TCACTTTGGT ACCAGGCTAC ATGGTGGCAA GGATTCTGCT AGTCCACGAT 2520
ACATCTTTAC AATGCTCAGC TCTTTGGCTC GATGTTATT TCCACAAA GATGATCACA 2580
CGTTGAAGTT TTTATATGAT GACAAACGAG GTGTGAGCC TGAATGGTAC ATTCTTATTA 2640
TTCCCATGGT GCTGATAAAT GGTGCTGAAG GAATCGGTAC TGGGTGGTCC TGCAAAATCC 2700
CCAACTTTGA TGTGCTGAAG ATTGTAAATA ACATCAGGCG TTTGATGGAT GGAGAAGAAC 2760
CTTTGCCAAT GCTTCCAAGT TACAAGAACT TCAAGGGTAC TATTGAAGAA CTGGCTCCAA 2820
ATCAATATGT GATTAGTGGT GAAGTAGCTA TTCTTAATTC TACAACCAT GAAATCTCAG 2880
AGCTTCCCGT CAGAACATGG ACCCAGACAT ACAAAGAACA AGTTCTAGAA CCCATGTTGA 2940
ATGGCACCAG GAAGACACCT CCTCTCATAA CAGACTATAG GGAATACCAT ACAGATACCA 3000
CTGTGAAAT TGTGTGAAG ATGACTGAAG AAAAAGTGGC AGAGGCAGAG AGAGTTGGAC 3060
TACACAAAGT CTTCAAACCT CAACTAGTC TCACATGCAA CTCTATGGTG CTTTGTGACC 3120
ACGTAGGCTG TTTAAAGAAA TATGACACGG TGTGGSATAT TCTAAGAGAC TTTTGTGAAC 3180
TCAGACTTAA ATATTATGGA TTAAGAAAAG AATGGCTCTT AGGAATGCTT GGTGCTGAAT 3240
CTGCTAAACT GAATAATCAG GCTCGCTTTA TCTTAGAGAA AATAGATGGC AAAATAATCA 3300

5
 10
 15
 20
 25
 30
 35
 40

```

TTGAAATAA GCCTAAGAAA GAATTAATTA AAGTTCCTGAT TCAGAGGGGA TATGATTCCG 3360
ATCCTGTGAA GGCCTTGAAA GAAGCCAGC AAAAGGTTCC AGATGAAGAA GAAATGAAG 3420
AGAGTGACAA CGAAAGAGAA ACTGAAAAGA GTGACTCCGT AACAGATTCT GGACCAACCT 3480
TCAACTATCT TCTTGATATG CCCCTTGGT ATTTAAACCA GGAAGAGAAA GATGAATCTCT 3540
GCAGGCTAAG AAATGAAAAA GAACAAGAGC TGGACACATT AAAAGAGAA AGTCCATCAG 3600
ATTGTGTGAA AGAAGACTTG GCTACATTTA TTGAAGAATT GGAGGCTGTT GAAGCCRAAG 3660
AAAAACAAGA TGAACAAGTC GGACTTCCTG GGAAGGGGGG GAAGGCCAAG GGGAAAAAAA 3720
CACAAATGGC TGAAGTTTTC CCTTCTCCGC GTGGTCAAG AGTCATTCCA CGAATAACCA 3780
TAGAAATGAA AGCAGAGGCA GAAAAGAAAA ATAAAAAGAA AATTAAGAAT GAAATACTG 3840
AAGGAAGCCC TCAAGAAGAT GGTGTGGAAC TAGAAGGCCT AAAACAAGA TTAGAAAAGA 3900
AACAGAAAAG AGAACCAGGT ACAAGAGCAA AGAAACAAAC TACATTGGCA TTTAAGCCAA 3960
TCAAAAAAGG AAAGAAGAGA AATCCCTGGC CTGATTGAGA ATCAGATAGG AGCAGTGACG 4020
AAAGTAATTT TGATGTCCCT CCACGAGAAA CAGAGCCAGC GAGAGCAGCA ACAAACAA 4080
AATTCACAA TGAATTGGAT TCAGATGAAG ATTTCTCAGA TTTTGTAGAA AAAACTGATG 4140
ATGAAGATTT TGTCCTATCA GATGCTAGTC CACCTAAGAC CAAAACCTCC CAAAACCTTA 4200
GTAACAAAGA ACTGAAACCA CAGAAAAGTG TCGTGTGAGA CCTTGAAGCT GATGATGTTA 4260
AGGGCAGTGT ACCACTGTCT TCAAGCCCTC CTGCTACACA TTTCCAGAT GAAACTGAAA 4320
TTACAAACCC AGTTCCTAAA AAGAAATGTA CAGTGAAGAA GACAGCAGCA AAAAGTCAGT 4380
CTTCCACCTC CACTACCGGT GCCAAAAAAA GGGCTGCCCT AAAAGGAACT AAAAGGGATC 4440
CAGCTTTGAA TTTCTGGTGC TCTCAAAAGC CTGATCTGCG CAAAACCAAG AATGCCGCA 4500
AAAGGAAGCC ATCCACTTCT GATGATTCTG ACTCTAATTT TGAGAAAATT GTTTCGAAAG 4560
CAGTCAAGG CAAGAAATCC AAGGGGGAGA GTGATGACTT CCATATGGAC TTTGACTCAG 4620
CTGTGGCTCC TCGGGCAAAA TCTGTACGGG CAAAGAAACC TATAAGTAC CTGGAAGAGT 4680
CAGATGAAGA GATCTGTGTT TAAATGTGA GCGGATTATT TTAAGTAATT ATCTTACCAA 4740
GCCCAAGACT GGTTTTAAAG TTACCTGAAG CTCTTAACTT CCTCCCTCT GAATTTAGTT 4800
TGGGAAGGT GTTTTATGTA CAAGACATCA AAGTGAAGTA AAGCCCAAGT GTTCTTAGC 4860
TTTTTAAAT ACTGCTTAAA TAGTGACCAT CTCATGGGCA TTGTTTCTT CTCTGCTTG 4920
TCGTGTTTT GAGTCTGCTT TCTTTGTCT TTAACCTG ATTTTAAAGT TCTTCTGAAC 4980
TGTAAGAAAT GCTATCTGAT CACTTCAGC TAAAGCAGTG TGTATTATA CCATCCAATA 5040
AGCTAAACT AGAGCAGTTT GATTTAAAG TGTCACCTT CCTCCTTTC TACTTTCAGT 5100
AGATATGAGA TAGAGCATAA TTATCTGTTT TATCTTAGTT TTATACATAA TTTACCATCA 5160
GATAGAACT TAGGTTCTTA GTACAGATAC TCTACTACAC TCAGCCTCTT ATGTGCCAAG 5220
TTTTCTTTA AGCAATGAGA AATTGCTCAT GTTCTTCATC TTCTCAAATC ATCAGAGGCC 5280
AAGAAAAAC ACTTTGGCTG TGTCTATAAC TTGACACAGT CAATAGAAATG AAGAAATTA 5340
GAGTAGTTAT GTGATTATTT CAGCTCTTGA CCTGTCCCTT GTGGTGCCT CTGAGTCTGA 5400
ATCTCCCAAA GAGAGAAACC AATTCTTAAG AGGACTGGAT TGCAGAGAC TCGGGGACAA 5460
CATTTGATCC AAGATCTTAA ATGTTATATT GATAACCATG CTCAGCAATG AGCTATTAGA 5520
TTCAATTTGG GAAATCTCCA TAAATTCAT TTGTAAACTT TGTTAAGACC TGCTACATTT 5580
GTTATATGTT GTGACTTGA GTAATGTTAT CAACGTTTTT GTAAATATTT ACTATGTTTT 5640
TCTATTAGCT AAATTCACAA AATTTGTGAC TTTAATAAAA TGTTCTAAAC ATTGC
  
```

Seq ID NO: 178 Protein sequence:
 Protein Accession #: NP_001058.1

45
 50
 55
 60
 65
 70

```

1 11 21 31 41 51
MEVSPLPVNV ENMQVNVKIK NEDAKKRLSV ERIYQKKTQL EHILLRPDPT IGSVELVTQQ 60
MMVYDEVDGI NYREVTFVPG LYKIFDEILV NAADNRQRDP KMSCIRVTID PENNLISIWN 120
NGKGFVVEH KVEKMYVPAI IFGQLLTSSN YDDEKKVTG GRNGYGAKLC NIPSTKFTVE 180
TASREYKMF KQTWMDNMGR AGEMELKPFM GEDYTCTIFQ PDLSEKFMQS LDKDIALMV 240
RRAYDIAGT KDVKVFLNGN KLPVKGFRSY VDMYLDKDLK ETGNSLKVH EBNVHRWEVC 300
LTMSEKPGQ ISFVNSIATS KGGRHVDYVA DQIVTKLVDV VKKKNKGGA VKAHQVKNHM 360
WIFVALIEN PTFDSQTKEN MTLQPKSPGS TCQLSEKPIK AATGCGIVES ILNWVFKPAQ 420
VQLNKKCSAV KENRIKIPK LDDANDAGGR NSTECTLILT BGDSAKTLAV SGLGVVGRDK 480
YGVFPLRGKI LNVREASHKQ IMENAEENNI IKIVGLQYKK NYEDEDLKT LRYGKIMINT 540
DQDQDGHSHK GLLINFHNH WPSLLRHRFL EEFITPIVKV SKNKQEMAFY SLPEFEWKS 600
STENHKWKV KYKGLGTST SKEAKEYPAD MKRHRIQPKY SGPEDDAIS LAFSKQIDD 660
RKEWLTFME DRQRKLGLL PEDYLYGQTT TYLTYNDFIN KELILFNSD NERSIPSMVD 720
GLKPGQRVVL FTCPKRNDRK EVKVAQLAGS VAEEMSSYHG EMSLMTIIN LAQNFVGSNN 780
LNLQPIQGF GTRLHGKDS ASPRYFTML SSLARLLFPF KDDHTLKFLY DDNRVPEW 840
YIPIPMVLI NGAGIGTGW SCKIPNFVDR EIVNNIRRLM DGEPLPMLP SYKNFKGTIE 900
ELAPNQYVIS GEVAILNSTT IEISELPVRT WTQTYKEQVL EPLNGTEKT PPLITDYREY 960
HTDTTVKPVV KMTBEKLAEA ERVGLHKVFK LQTSLTCSNM VLFDHVGLCK KYDTVLDIR 1020
DPFELRLKY GLRKEWLLGM LGAESAKLNN QARFILEKID GKIIENKPK KELIKVLIQR 1080
GYSDPVRKAW KEAQKQVPEE EENEESDNKE ETEKSDSVTD SGPTFNYLLD MPLWYLTKEK 1140
KDELCLRLNE KEQELDTLKR KSPSLWKEK LATFIELEA VEAKKQDEQ VGLPGKGGKA 1200
KKGKTQMAEV LPSPRGRVI PRITIEKAE AEKKNKKIK NENTEGSPQ DGVELEGLKQ 1260
RLEKKQKREP GTTKKQTTL AFKPIKGGK RNPWPDESSED RSSDESNDV PPRETEPRRA 1320
ATKTFTMDL DSEDEFSDP EKTDDDPVP SDASPPKTKT SPKLSNKLK PQKSVVSDLE 1380
ADDVKGSVPL SSSPPATHTP DETEITNPVP KKNVTVKKTA AKSQSSTST GAKKRAAPKG 1440
TKRDPALNSG VSQKPDPAKT KNRRKRKPST SDDSDSNFEK IVSKAVTSKK SKGESDDPHM 1500
DFDSAVAPRA KSVRAKPIK YLEESDEDDL F
  
```

Seq ID NO: 179 DNA sequence
 Nucleic Acid Accession #: E08 sequence
 Coding sequence: 148-7095

80
 85

```

1 11 21 31 41 51
CACACATACG CACGCACGAT CTCACTTCCA TCTATACACT GGAGGATTAA AACAAACAAA 60
CAAAAAAACC ATTTCCCTCG CTCCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120
CGCGAGGGGG CGCGAGACCG TCTGAAAATG CGAATCTTAA AGCGTTTCTC CGCTTGCAAT 180
CAGCTCCTCT GTGTTTGCCG CTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
CTTGTGAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAA TTGGGGAAAG 300
AATATACCAA CATGTAATAG CCAAAAAACA TCTCCTATCA ATATTGATGA AGATCTTACA 360
CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTACGGGTT GGGATAAAAC ATCATGGAA 420
  
```

	AACACATTCA	ITCATAACAC	TGGGAAACA	GTGGAAATTA	ATCTCATAA	TGACTACCGT	489
	GTGAGGAG	GAGTTTCAGA	AATGGTGT	AAAGCAAGCA	AGATAACTTT	TCAGTGGGA	540
	AAATGCAATA	TGTATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAA	660
5	GGAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTGAGG	TTGGACAGA	AGAAAATTG	720
	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
10	ACAGTTAGCA	TCTCTGAAG	CCAGTTGGCT	GTTTTGTG	AAGTCTTAC	AATGCAACAA	960
	TCTGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTGGAGAGCA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTT	CTCATACACT	GGAAAGGAAG	AGATTTCATGA	AGCAGTTTGT	1080
	AGTTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTGTTTACA	1140
	TGGGAAAGAC	TGCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGAGT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
15	GGTGCTATT	TCATAAATT	GCTACCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TGATCTTTT	COCTGAATTA	ATTGGAAGT	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGA	AAGACATGA	AGAAGGCGCT	ATTGTGAATC	CTGTAGAGA	CAGTGCTACA	1500
20	AACCAATCA	GGAAAAGGA	ACCCAGATT	TCTACCAAA	CACACTACAA	TGCGATAGGG	1560
	ACGAATACA	ATGAAGCCA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACAGTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCTCT	GACTTCTCAG	ACTGTGACTG	AACCTGCCAC	TCACACTGTG	1740
	GAAGGTACTT	CAGCTCTCTT	AAATGATGGC	TCTAAAACCT	TTCTTAGATC	TCCCATATG	1800
25	AACCTGTGCG	GGAGTCGAGA	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGTATAT	ATTTCTCTCC	1980
	GAAAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCACTAGG	TTCAGAAGAA	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAT	2100
30	GTGTGTTTTT	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACGTAGATA	CGTGTGATG	AATCTGAGAA	GACAAACCA	2220
	TCCCTTTCTG	CAGGCCAGT	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTTCTA	CCTTTGCTTA	CTTCCCAACT	GAGGTAAAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTTGGT	CTCCACGGTC	AACGTGGTAT	ACTGCGAGAC	AACCCCAACG	2400
35	GTATACAATG	GTGAGACACC	TCTTCAACCT	TCCTACAGTA	GTGAAGTCTT	TCCTCTAGTC	2460
	ACCCCTTTGT	TGCTTGACAA	TCAGATCCTC	AACACTACCC	CTGCTGCTTC	AAGTAGTGAT	2520
	TCCGCTTTGC	ATGCTACGCC	TGTATTTCCC	AGTGTGATG	TGTCAATTGA	ATCCATCTCT	2580
	TCTTCTCTAT	ATGCTGACCC	TTTGCTTCCA	TTTCTCTCTG	CTTCTCTCAG	TAGTGAATTG	2640
40	TTTGCCTATC	TGCATACAGT	TTCTCAAACT	CTTCCACAAG	TTACTTCAGC	TACCGAGAGT	2700
	GATAAGTGCT	CCTTGATGCT	TTTCTGCCA	GTGGCTGGGG	GTGATTGCT	ATTAGAGCCC	2760
	AGCCTTGCTG	AGTATCTGTA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GACGCTGGAA	2820
	TTTGGTAGTG	AATCTGGTGT	TCTTTATAAA	ACGCTTATGT	TTTCTCAAGT	TGAACCAACC	2880
	AGCAGTGATG	CCATGATGCA	TGCACGTTCT	TCAGGGCCTG	AACCTTCTTA	TGCTTGTCT	2940
45	GATAAAGGAG	GCTCCCAACA	CATCTTCACT	GTTCCTTACA	GTCTGCAAT	ACCTGTGCA	3000
	GATTCTGTGG	GTGTAACCTA	TCAGGGTTCC	TTATTTAGCG	GCCCTAGCCA	TATACCAATA	3060
	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGATGGGG	AATGGTCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAAACAT	TTCTTCACT	GTTCCTGTAG	CTGAATTTAC	ATATACAACA	3240
50	TCTGTGTTTG	GTGATGATAA	TAAGGCGCTT	TCTAAAAGTG	AAATAATATA	TGAAAATGAG	3300
	ACTGAAGTGC	AAATTCCTTC	TTCAATGAG	ATGGTTTACC	CTTCTGAAG	CACAGTCACT	3360
	CCCAACATGT	ATGATAATGT	AAATAAGTTG	AATGCGTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTTCTAGCA	CCAAGGGCAT	GTTTCCAGGG	TCCCTTGCTC	ATACCACCAC	TAAGGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAAT	AACTTTTTCA	TTCAACCTAC	ACATACGTGC	3540
55	TCTCAAGCAT	CTGGTGACAC	TTGCTTAAAC	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
	TCCTCTGACC	CTGCTTCTAG	TGAATATGTA	TCTCCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
	ACCTCAGCTT	CGTTAATAAC	TGAAGTATG	CTACAACCTT	CCTTTACAGG	TTCTGATGTT	3720
	GACACCTTGC	TAAAACTGT	TCTTCCAGCT	GTGCCACAGT	ATCCAATATT	GGTGAAACCC	3780
	CCCAAGTTGT	ATGAATTTAG	TTCTACAATG	TTGCATCTCA	TTGTATCMAA	TTCTGCTTCA	3840
60	AGTGAAAACA	TGCTGCACTC	TACATCTGTA	CCAGTTTTTG	ATGTGTGCGC	TACTTCTCAT	3900
	ATGCACCTCT	CTTCACTTCA	AGGTTTGACC	ATTTCTCTAT	CAAGTGAGAA	ATATGAACCA	3960
	GTTTTGTAAA	AAAGTGAAAG	TTCCACCAA	GTGGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCAAA	CGGCCAATTT	GGAGATTAA	CAGGCCCATC	CCCCAAAAGS	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTTATC	AATTGATGAA	CCATTAAATA	CACTAATAAA	TAAGCTTATA	4140
65	CATTCCGATG	AAATTTTAAAC	CTCCACCAA	AGTTCTGTTA	CTGGTAAGGT	ATTTGCTGGT	4200
	ATTCCAACAG	TTGCTTCTGA	TACATTTGTA	TCTACTGATC	ATTCGTGTTCC	TATAGGAAAT	4260
	GGGCATGTTG	CCATTACAGC	TGTTTCTCCC	CACAGAGATG	GTTCGTAAAC	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACCTTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCCGGT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGATGAC	4440
70	AGAGGTAGTG	ATGGCTTATC	CATTCTAAG	TGTATGTCAT	GCTCATCTTA	TAGAGAAATCA	4500
	CAGGAAAAGG	TAATGAATGA	TTCAGACACC	CACGAAAACA	GTCTTATGGA	TCAGAAATAAT	4560
	CCAACTCTCAT	ACTCACTATC	TGAGAATTCT	GAAGAAGATA	ATAGAGTCAC	AAGTGATATCC	4620
	TCAGACAGTC	AAACTGGTAT	GGACAGAGT	CCTGGTAAAT	CACCATCAGC	AAATGGGCTA	4680
	TCCCAAAAGC	ACATGATGG	AAAAGAGGAA	AATGACATTC	AGACTGGTAG	TGCTCTGCTT	4740
75	CCTCTCAGCC	CTGAATCTAA	AGCATGGGCA	GTTCYGACAA	GTGATGAAGA	AAGTGGATCA	4800
	GGGCAAGGTA	CCTCAGATAG	CCTTAATGAG	AATGAGACTT	CCACAGATT	CAGTTTGTCA	4860
	GACACTAATG	AAAAGATGC	TGATGGGATC	CTGGCAGCAG	GTGACTCAGA	AATAACTCCT	4920
	GGATTCCAC	AGTCCCCAAC	ATCATCTGTT	ACTAGCGAGA	ACTCAGAAGT	GTTCACGTT	4980
	TCAGAGGAG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTGTA	TTGGTCTAGC	TGAGGGGTTG	5040
80	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTATCTGTT	5100
	CTAGTGGTTC	TTGTTGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGACACATTT	5160
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	5220
	ATTTCAAGTG	ATGTCGGAGC	AATTCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	5280
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTTT	CCAGGAAGTG	5340
	CAGAGCTGTA	CTGTTGACTT	AGGTATTACA	GCAGACAGCT	CCAACCAACC	AGACAAACAG	5400
85	CACAGAAATC	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	5460
	CTTGCTGAAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTTATG	TGATGGCTAC	5520
	AACAGACCAA	AAGCTTATAT	TGCTGCCCAA	AGCCCACTGA	AATCCACAGC	TGAAGATTTT	5580
	TGGAGAATGA	TATGGGAACA	TAAATGGGAA	GTTATTGTCA	TGATAACAAA	CCTGTGGAG	5640

5
10
15
20
25
30
35
40

AAAGGAAGGA GAAATGTGA TCAGTACTGG CCTGCCGATG GGAGTGGAGGA GTACGGGAAC 5700
 TTTCTGGTCA CTCAGAAGAG TGTGCAAGTG CTTCGCTATT ATACTGTGAG GAAITTTTACT 5760
 CTAAGAAACA CAAAATAAAA AAAGGGCTCC CAGAAAGGAA GACCCAGTGG ACCTGTGGTC 5820
 ACACAGTATC ACTACACGCA GTGGCCTGAC ATGGGAGTAC CAGAGTACTC CCTGCCAGTG 5880
 CTGACCTTTG TGAGAAAGGC AGCCTATGCC AAGCGCCATG CAGTGGGGCC TGTGTGTGTC 5940
 CACTGCAGTG CTGGAGTTGG AAGAACAGGC ACATATATTG TGCTAGACAG TATGTTGCAG 6000
 CAGATTCAAC ACGAAGGAAC TGTCAACATA TTTGSCITCT TAAACACAT COGTTTCAAA 6060
 AGAAATTATT TGGTACAAAC TGAGGAGCAA TATGCTTCA TTCATGATAC ACTGGTTGAG 6120
 GCCATACCTA GTAAAGAAAC TGAGGTGCTG GACAGTACATA TTCATGCCCTA TGTTAATGCA 6180
 CAGTCAATA TACAGCAGAG TGACTATTCT GCAGCCCTAA AGCAATGCAA CAGGGAAGAG 6240
 AATGGAATT CTTCATCAT CCTGTGGAA AGATCAAGGG TTGGCATTTC ATCCTGAGT 6300
 GGAGAAGGCA CAGACTACAT CAATGCTCTC TATATCATGG GCTATTACCA GAGCAATGAA 6360
 TTCATCATA CCCAGCACCC TCTCCTTCAT ACCATCAAGG ATTTCTGGAG GATGATATGG 6420
 GACCATAAAG GGCACCTGGT GGTATGATT CCTGATGGCC AAAACATGGC AGAAGATGAA 6480
 TTTGTTTACT GGCCAAATAA AGATGAGCCT ATAAATTGTG AGAGCTTTAA GGTCACTCTT 6540
 ATGGCTGAG AACACAAATG TCTATCTAAT GAGGAAAAC TTATAATTCA GGACTTTATC 6600
 TTAGAAGCTA CACAGGATGA TTATGTAATT GAAGTGAGGC ACTTTCAGTG TCCTAAATGG 6660
 CCAATCCAG ATAGCCCATC TAGTAAACTT TTTGAACCTA TAAGTGTAT AAAAGAAGAA 6720
 CTGCGCAATA GGGATGGGCG TATGATTGTT CATGATGAGC ATGGAGGAGT GACGGCAGAA 6780
 ACTTCTGTG CTCTGACAC CCTTATGCAC CAATAGAAA AAGAAATTC CGTGGATGTT 6840
 TACCAGGTAG CCAAGATGAT CAATCTGATG AGGCCAGGAG TCTTGTCTGA CATTGAGCAG 6900
 TATCAGTTTC TCTACAAAGT GATCCTCAGC CTGTGAGCA CAAGGCAGGA AGAATGCCA 7020
 TCCACCTCTC TGGACAGTAA TGGTGCAGCA TTGCTGATG GAAATATAGC TGAGAGCTTA 7080
 GAGTCTTAG TTTAAACGAG AAAGGGGTGG GGGGACTCAC ATCTGAGCAT TGTTTTCTCT 7140
 TTCTAAATAG TAGGCAGGAA AATCAGTCTA GTTCTGTAT CTGTTGATTT CCATCACTCT 7200
 GACAGTAATC TTCATGACAT AGGATTCTGC CGCCAAATTT ATATCATTAA CAATGTGTGC 7260
 CTTTTTGCAA GACTTGTAA TTACTTATTA TGTTTGAAT AAAATGATTG AATTTTACAG 7320
 TATTTCTAAG AATGGAATG TGGTATTTT TTTCTGATTG ATTTTAACAG AAAATTTCAA 7380
 TTTATAGAG TTAGGAATG CAAACTACAG AAAATGTTTG TTTTATAGTG CAAATTTTAA 7440
 GCTGTATTTG TAGCAATAT CAGGTTTGT AGAAATATAA CTTTAAATAC AGTAGCCTGT 7500
 AAATAAAACA CTCTCCATA TGATATTCAA CATTTTACAA CTGAGTATT CACCTAAAGT 7560
 AGAAATAATC TGTACTTAT TGTAAATACT GCCCTAGTGT CTCCATGGAC CAAATTTATA 7620
 TTTATAATG TAGATTTTAT TATTTTACTA CTGAGTCAAG TTTTCTAGTT CTGTGTAATT 7680
 GTTAGTTTA ATGACGTAGT TCATTAGCTG GTCTTACTCT ACCAGTTTTC TGACATTGTA 7740
 TTGTGTTACC TAAGTCATTA ACTTTGTTTC AGCATGTAAT TTTAACTTT GTGGAATAA 7800
 GAAATACCTT CATTTTGAAG GAAGTTTATA TGAGAATAAC ACCTTACCAA ACATTTGTTCA 7860
 AATGGTTTTT ATCCAAGGAA TTGCAAAAT AAATATAAAT ATTGCCATTA AAAAAAATA 7920
 AAAAAAAAAA AAAAAAAAAA AAAA

Seq ID NO: 180 Protein sequence:
 Protein Accession #: Eos sequence

45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
 MRILKRLAC IQLLCVCRLL WANGYYRQR KLVEEIGWSY TGALEQKNGW KKPTCNSPK 60
 QSPINIDEDL TQVNVNLKLL KFGQWDKTSL ENTFIHNTGK TVEINLINDY RVSGGVSEMV 120
 FKASKITFWH GKCMSSSDGS EHSLEGGKFF LEMQIYCFDA DRFSSFEEAV KKGKGLRLALS 180
 ILFEVGTEN LDFKAITDGV ESVSRFGKQA ALDPFILLNL LPNSTDKYIY YNGSLTSPPC 240
 TDTVDWIVFK DTVSISESQL AVFCEVLTMQ QSGVLMMDY LQNNFRQQY KFSRQVFSY 300
 TGKEEIEHAV CSSSEPIVQA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360
 HEFLTDGYQD LGAILNLNLL NMSYVLQIVA ICTNGLYKGY SDOLIVDMPT DNPDLDFPE 420
 LIGTEEIKKE EEEGKDIEEG AIVNPGRDSA TNQIRKKEPO ISTTTHYMRI GTKYNSAKTN 480
 RSPTRGSEFS GKGDVNPTSL NSTSQPVTKL ATEKDLSLTS QVTTELPFHT VEGTSASLND 540
 GSKTVLRSFH MKLSGTAESL NTVSITEYEE ESLTSFKLD TGAEDSSGSS PATSAIPFIS 600
 ENISQGVIPS SENPETITVD VLIPEASRNA SEDSTSSGSE ESLKDPSEMG NWFPSSTDI 660
 TAQPDVSGSR ESPLQNTYTE IRVDESEKTT KSFAGPVMS QGSPVTDLEM PHYSTPAYFP 720
 TEVTPHAFPT SSRQDLVYST VNVVYSQTTO PVYNGETPLQ PSYSSEVFPL VTPLLLDNQI 780
 LNTTPAASSS DSALHATPVF PSVDVSPESI LSSYDGLAPL PFSSASPSSE LFRHLHTVSQ 840
 ILPQVTATE SDKVPLEASL PVAGGDLLE PSLAQYSDVL STTHAASETL EFGSESGVLY 900
 KTLMPQVPEP PSSDAMMHAR SSGPEPSYAL SDNEGSQHIF TVSYSSAIPV HDSVGVTYQG 960
 SLFSGPSHIP IPKSLITPT ASLLQPTHAL SGDGEWSGAS SDSEPLLPDT DGLTALNISS 1020
 PVSVAEFTYF TSVFGDDNKA LSKSEIYGN ETELQIPSFN EMVVPSESTV MPNMYDNVVK 1080
 LNASLQETSV SISSTKGMPF GSLAHTTTKV FDHEISQVPE NNFVSQPTHV VSQASGDTSL 1140
 KPVLSANSEP ASSDPASSEH LSPSTQLLFY ETSASFSTEV LLQPSFQASD VDTLLKTVP 1200
 AVPSDPILVE TPKVDKISST MLHLIVNSA SSENMLHSTS VPVFDVSPTS HMHSASLQGL 1260
 TISYASEKYE PVLLKSESSH QVPSLYSND ELFQTANLEI NQAHPKGRH VFATFVLSD 1320
 EPLNTLINKL IHSDEILTST KSSVTGKVF A GIPTVASDTE VSTDHVPF IG NGRVAITAVS 1380
 PHRDGVSST KLLFPKATS ELSHSAKSDA GLVGGGEGD DDDGDDDDDD DRGSDGLSIH 1440
 KCMSCSSYRE SQEKVMNDS THENSMDQN NPISYLSSEN SEEDNRVTSV SSDSQTGMDR 1500
 SPGKSPSANG LSQKHNDGKE ENDIQTSAL LPLSPESKAW AVLTSDSEESG SQQGTSDSLN 1560
 ENETSTDFSP ADTNEKDADG ILAAGDSEIT PGFPQSPSTSS VTSENSEVPH VSEAEASNSS 1620
 HESRIGLAEG LESEKKAVIP LVIVSALTPI CLVVLVGILI YWRKCPQTAH FYLEDSTSPR 1680
 VISTPPTPIF PISDDVGAIP IKHPPKHVAD LHASSGFTEE FETLKEFYQE VQSCVTDLGI 1740
 TADSSNHPDN KHKRYINIV AYDHSRVKLA QLAEKDGKLT DYINANYVDG YNRKAYIAA 1800
 QGFLKSTAEF FMRMIWEHNV EVIVMITNLV EKRGRKCDQY WPDGSEEEY NFLVTQKSVQ 1860
 VLAYYTVRNP TLRNTKIKKG SQKGRPSGRV VTQYHYTQWP DMGVPEYSLP VLTFFVRKAA 1920
 AKRHAVGPVV VHCAGVGRV GTYIVLDSML QQIQHEGTVN IFGFLKHRS QRNVLVQTEE 1980
 QVVPIDHTLV EAILSKETEY LDSHIHAYVN ALLIPGPAGK TKLEKQFOLL SQSNIQSDY 2040
 SAALKQCNRE KNRTSIIIPV ERSRVGISSL SGEQTDYINA SYIMGYQSN EPIITQHPLL 2100
 HTIKDFWRMI WDHNAQLVVM IPDQONMAED EFVWPNKDE PINCESPKVT LMAEHLKCLS 2160
 NEEKLIQDF ILEATQDDYV LEVRHFCQPK WPNPDSPISK TFEELISVIK EAANDRGPMI 2220
 VDEHGGVTA GTFCALTTLA HQLEKENSVD VYQVAKMINL MRPGVFADIE QYQFLYKVL 2280
 SLVSTRQSEN PSTSLDSNGA ALPDGNIAES LESLV

Seq ID NO: 181 DNA sequence
 Nucleic Acid Accession #: Eos sequence

	1	11	21	31	41	51	
5	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGCGAGGGG	CGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTCCCG	CCTGGATTGG	GCTAATGSGT	ACTACAGACA	ACAGAGAAAA	240
10	CTTGTTGAAG	AGATTGGCTG	GTCTATACA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAACCTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTTGGA	420
	AACACATTCA	TTCTATAACAC	TGGGAAACAA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTCAGCGGAG	GAGTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
15	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAAATTG	720
	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACGTGT	GAACCTTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTTAC	840
20	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAG	CCAGTTGGCT	GTTTTGTGTC	AAGTTCTTAC	AATGCAACAA	960
	TCTGGTTATG	TCTGCTGATG	GGACTACTTA	CAAAACAATT	TTGAGAGACA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTCTATG	AGCAGTTTGT	1080
	AGTTCAGAAC	CAGAAAAATG	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTACA	1140
25	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGCACT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTC	TCAATTAATT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	CTCTTATGGS	AAAATACAGC	GACCAACTGA	TTGTCGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAACCT	AAGAAATAAT	CAAGAGGAGG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGGGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
30	AACCAATACA	GGAAAAAGGA	ACCCACAGAT	TCTACCAACA	CACACTACAA	TGCGATAGGG	1560
	ACGAATAACA	ATGAAGCCAA	GACTAACCCA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTAGTG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCCTT	GACTTCTCAG	ACTGTGACTG	AACGCCACCC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAACCT	TTCTTAGATC	TCCACATATG	1800
35	AACCTTGTGG	GGACTGCGAG	ATCCTTAAAT	ACAGTTTCTA	TAAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTTA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTCTCT	GCTTCCCATG	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAGAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
40	GAAGATTCAA	CTTCATCAGG	TTCAGAAGAA	TCATAAAGG	ATCCTTCTAT	GGAGGSAAT	2100
	GTGTGTTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTGATG	AATCTGAGAA	GACAAACCAAG	2220
	TCTTTTCTCT	CAGGCCAGTG	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CCTTTGCCCTA	CTTCCCAACT	GAGGTAAACAC	CTCATGCTTT	TACCCCATCC	2340
45	TCCAGACAAC	AGGATTGGGT	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCCAACG	2400
	GTATACAATG	CAGAGGCCAG	TAATAGTAGC	CATGAGTCTC	GTATTGGTCT	AGCTGAGGGG	2460
	TTGGAATCCG	AGAAAGAGG	AGTTATACCC	CTTGATGATG	TGTCAGCCCT	GACTTTTATC	2520
	TGTTAGTCTG	TCTTGTGGGG	TATTCTCATC	TACTGGAGGA	AATGCTTCCA	GACTGCACAC	2580
	TTTTACTTAG	AGGACAGTAC	ATCCCTTAGA	GTTATATCCA	CACCTCCAAC	ACCTATCTTT	2640
50	CCAATTTCAG	ATGATGTCCG	AGCAATTTCA	ATAAAGCACT	TTCCAAAGCA	TGTTGCAGAT	2700
	TTACATGCAA	GTAGTGGGTT	TACTGAAGAA	TTTGAGACAC	TGAAGAGATT	TTACAGGAA	2760
	GTGCAGAGCT	GTACTGTTGA	CTTAGGTATT	ACACAGAGCA	GCTCCAACCA	CCAGACAAC	2820
	AAGCACAAGA	GTACTATACAT	AAATATCGTT	GCCTATGATC	ATAGCAGGGT	TAAGCTAGCA	2880
	CAGCTTGCTG	AAAAGGATGG	CAACTGACT	GATTATATCA	ATGCCAATTA	TGTTGATGGC	2940
55	TACAAAGAGC	CAAAAGCTTA	TATTGCTGCC	CAAGGCCAC	TCAAAATCCAC	AGCTGAAGAT	3000
	TTCTGGAGAA	TGATATGGGA	ACATAATGTG	GAAGTTATTG	TCATGATAAC	AAACCTCGTG	3060
	GAGAAAGGAA	GGAGAAATG	TGATCAGTAC	TGGCCTGCCG	ATGGGAGTGA	GGAGTACGGG	3120
	AACCTTCTGG	TACTCAGAAA	GAGTGTGCAA	GTGCTTGCCCT	ATTATACGTG	GAGGAATTTT	3180
	ACTCTAAGAA	ACACAAATAT	AAAAAGGGGC	TCCAGAGAA	GAAGACCCAG	TGGACGTGTG	3240
60	GTCCACAGCT	ATCACTACAC	GCACTGGCCT	GACATGGGAG	TACCAGAGTA	CTCCCTGCCA	3300
	GTGCTGACCT	TTGTGAGAAA	GGCAGCCTAT	GCCAGAGGCC	ATGCACTGGG	GCCTGTGTGC	3360
	GTCCACTGCA	GTGCTGGAGT	TGGAGAACAA	GGCACATATA	TTGTGCTAGA	CAGTATGTTG	3420
	CAGCAGATTG	AACACGAAAG	AACCTGTCAAC	ATATTGGGCT	TCTTAAACAA	CATCCGTTCA	3480
	CAAGAAATTT	ATTTGTGACA	AACGTAGGAG	CAATATGTCT	TCATTATATG	TACACTGGTT	3540
65	GAGGCCATAC	TAGTAAAGA	AACGTAGGAG	CTGGACAGTC	ATATTCTATG	CTATGTTAAT	3600
	GCACCTCTCA	TTCTGGAGCC	AGCAGGCAAA	ACAAAGCTAG	AGAAACAATT	CCAGCTCCTG	3660
	AGCCAGTCAA	ATATACAGCA	GAGTGACTAT	TCTGCAGCCC	TAAAGCAATG	CAACAGGGAA	3720
	AAGAAATCGAA	CTTCTCTAT	CATCCCTGTG	GAAAGATCAA	GGGTGGCAT	TTTATCCCTG	3780
	AGTGGAGAGG	GCACAGACTA	CATCAATGCC	TCCTATATCA	TGGGCTATTA	CCAGAGCAAT	3840
70	GAATTATCAT	TTACCCAGCA	CCCTCTCCTT	CATACCATCA	AGGATTTCTG	GAGGATGATA	3900
	TGGGACCATA	ATGCCCAACT	GGTGGTTATG	ATTCTTGATG	GCCAAAACAT	GGCAGAGAGT	3960
	GAATTTGTTT	ACTGGCCAAA	TAAAGATGAG	CCTATAAATT	GTGAGAGCTT	TAAGGTCACT	4020
	CTTATGGCTG	AAGAACACAA	ATGTCTATCT	AATGAGGAAA	AACCTATAAT	TCAGGACTTT	4080
	ATCTTAGAAG	CTACACAGGA	TGATTATGTA	CTTGAAGTGA	GGCACTTTCA	GTGCTCTAAA	4140
75	TGGCCAAATC	CAGATAGCCC	CATTAGTAAA	ACTTTTGAAC	TTATAAGTGT	TATAAAGAAA	4200
	GAAGCTGCCA	ATAGGGATGG	GCCTATGATT	GTTCATGATG	AGCATGGAGG	AGTGACGGCA	4260
	GGAACTTTCT	GTGCTCTGAC	AACCTTATG	CACCAACTAG	AAAAAGAAAA	TTCCGTGGAT	4320
	GTTTACCAAG	TAGCCAAAGT	GATCAATCTG	ATGAGGCCAG	GAGTCTTTGC	TGACATGTAG	4380
	CAGTATCAGT	TTCTCTACAA	AGTGATCCTC	AGCCTTGTTA	GCACAAGGCA	GGAGAGAGAT	4440
80	CCATCCACCT	CTCTGACAG	TAATGGTGCA	GCAATGGCCTG	ATGGAAATAT	AGCTGAGAGC	4500
	TTAGAGTCTT	TAGTTTAAAC	CAGAAAGGGG	TGGGGGAGCT	CACATCTGAG	CATTGTTTTC	4560
	CTCTTCTTAA	ACTTAGGCGA	GAAATTCAGT	CTAGTTCTGT	TATCTGTGTA	TTTCCCATCA	4620
	CCTGACAGTA	ACTTCTATGA	CATAGGATTC	TGCGGCCAAA	TTTATATCAT	TAACAATGTG	4680
	TGCCCTTTTG	CAAGACTTGT	AATTTACTTA	TTATGTTTGA	ACTAAATGA	TTGAATTTTA	4740
85	CAGTATTTCT	AAGAAATGAA	TTGTGGTAT	TTTTTCTGTA	TTGATTTTAA	CAGAAAAATT	4800
	CAATTTATAG	AGGTTAGGAA	TTCCAAACTA	CAGAAAAATG	TTGTTTTTAG	TGTCAAATTT	4860
	TTAGCTGTAT	TTGTAGCAAT	TATCAGGTTT	GCTAGAAATA	TAACCTTTAA	TACAGTAGCC	4920
	TGTAATAAAA	ACACTCTTCC	ATATGATATT	CAACATTTTA	CAACTGCAGT	ATTCACCTAA	4980

AGTAGAATA ATCTGTACT TATTGTAAT ACTGCCCTAG TGTCTCCATG GACCAAAATTT 5040
 ATATTATATA TTGTAGATT TTATATTTA CTACTGAGTC AAGTTTTCTA GTTCTGTGTA 5100
 ATTGTTTAGT TTAATGACGT AGTTCATTAG CTGCTCTTAC TCTACCAGTT TTCTGACATT 5160
 GTATTGTGTT ACCTAAGTCA TTAACCTTGT TTCAGCATGT AATTTTAACT TTTGTGGAAA 5220
 ATAGAAATAC CTTCATTTTG AAAGAAAGTTT TTATGAGAA AACACCTTAC CAAACATTGT 5280
 TCAATGGTT TTTATCCAAG GAATTGCAAA AATAAATATA AATATTGCCA TTAACAAAAA 5340
 AAAAAA AAAAAA AAAAAA

Seq ID NO: 182 Protein sequence:
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 MRILKRFLAC IQLLCVCRLLD WANGYYRQR KLVEEIGWSY TGLNQKNWG KKYPTCNSPK 60
 QSPINIDEDL TQVNVNKLKL KFGQWDTSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120
 FKASKITFHW GKCMSSDGS EHSLEGQKFP LEMQIYCFDA DRSSFEAEV KKGKRLRALS 180
 ILFEVGTEN: LDFKAIIDGV: ESVSFPGKQA ALDPPILLNL LPNSTDKYVI YNGSLTSPPC 240
 TDTVDIVFK DTVSISESQL AVFCEVLTMQ QSGYVLMMDY LQNNFREQQY KFSRQVFSSY 300
 TGKEEIEHAV CSSEPENVQA DPENYTSLLV TWERPRVVD TMIEKFAVLY QQLDGEDQTK 360
 HEPLTDGQD LGAILNLLP NMSYVLQIVA ICTNGLYKGY SDQLIVDMPT DNPELDLPPE 420
 LIGTEEIIKE EEEGKDIIEG AIVNPGRDSA TNQIRKKEPQ ISTTHYNRI GTKYNEAKTN 480
 RSPTRGSEFS KGGDVPNTSL NSTSQPVTKL ATEKDISLTS QTVTELPPT VEGTSASLND 540
 GSKTVLRSPH MNLGSGTAESL NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600
 ENISQGYIFS SENPETITVD VLIPEASRNA SEDSTSSGSE ESLKDPMEG NWFPSSDTI 660
 TAQPDVSGSR ESPLQNTYTE IRVDESEKTT KSFSAQPVMS QGSPVTDLEM PHYSTFAYFP 720
 TEVTPHAFIT SSRQDLVST VNVVYSQTTQ PVYNAAEASNS SHESRIGLAS GLESEKKAIV 780
 PLIVVSALT FICLVVLVGL IYWRKCFQTA HPYLEDSTSP RVISTPPTPI FPISDDVGAI 840
 PIKHPKHYVA DLHASSGFTE EFETLKEFYQ EVQSCVTDLG ITADSSNHDP NKHNRYINI 900
 VAYDHSRIVK AQLAEKDGKL TDYINANYVD GYNRPKAYIA AQGPLKSTAE DFWRMIEHN 960
 VEIVIMITNL VEGRRKCDQ YNPADGSEEV GNFLVTQKSV QVLAYYTVRN FTLRNTKIKK 1020
 GSQRGRPSGR VVTOYHYTQW PDMGVPEYSL PVLTFVRKAA YAKRHAVGPV VVHCAGVGR 1080
 TGYIYVLDLM LQIQHEGTV NIFGFLKHIR SQRYNLVQTE EQYVFIHDTL VEALSKETE 1140
 VLDSHIHAYV NALLIPGPAG RTKLEKQFQL LQSQNIQSD YSAALKQCNR EKNRTSSII 1200
 VERSRVGISS LSGEGTDYIN ASYIMGYQSS NEFIITQHPL LHTIKDPWRM IWDHNAQLVV 1260
 MIDPDQMAES DEFYWNKND EPINCESFKV TLMAEHKCL SNEEKLIQD FILEATQDDY 1320
 VLEVRHFQCP KWNPNPSPI S KTFELISVIK EEAANRDGPM IVHDEHGGVT AGTFCALTTL 1380
 MHQLEKENS DVYQVAKMIN LMRPGVFADI EQYQFLYKVI LSLVSTRQEE NPFSTLDSNG 1440
 AALFDGNIAE SLESIV

Seq ID NO: 183 DNA sequence
 Nucleic Acid Accession #: EOS sequence
 Coding sequence: 148-4494

1 11 21 31 41 51
 CACACATACG CAGCGACGAT CTCACCTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60
 CAAAAAAAC ATTTCTCTCG CTCCTCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120
 CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCCTAA AGCGTTTCTT CGCTTGCAAT 180
 CAGCTCCTCT GTGTTTGGCG CCGTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
 CTGTTGAAG AGATTGGCTG GTCCCTATACA GGAGCACTGA ATCAAAAAAA TTGGGGAAAG 300
 AATATCCAA CAGTAAATAG CCCAAACAAA TCTCCTATCA ATATTGATGA AGATCTTACA 360
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAAAC ATCATTTGAA 420
 AACACATTC TGCATAACAC TGGGAAAAAC GTGGAAATTA ATCTCACTAA TGACTACCGT 480
 GTCAGCGGAG GAGTTTCAGA AATGGTGTIT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540
 AATATCAATA TGCTACTGTA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600
 GAGATGCAAA TCTACTGCTT TGATGCAGAG CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660
 GGAAGAGGGA AGTTAAGAGC TTTATCCATT TTGTTGAGG TTGGGACAGA AGAAATTTTG 720
 GATTTCAAAG CGATTATGTA TGGAGTGGAA AGTGTTAGTC GTTTTGGGAA CGAGGCTGCT 780
 TTAGATCCAT TCATACTGTT GAACCTTCTG CCAACTCTCA CTGACAGTAA TTACATTAC 840
 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900
 ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTGTG AGTTCTTAC AATGCAACAA 960
 TCTGTTATG TCATGCTGAT GGACTACTTA CAAACAAAT TTCCAGAGCA ACAGTACAAG 1020
 TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAGGAAG AGATTCTAGA AGCAGTTTGT 1080
 AGTTTCAGAA CAGAAATATG TCAGGCTGAC CCAGAGAAAT ATACCAGCCT TCTTGTTACA 1140
 TGGAAAGAGC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGCAAT TTTGTACCA 1200
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260
 GGTGCTATTG TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320
 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTGCACAT GCCTACTGAT 1380
 AATCCTGAAC TTGATCTTTT CCCTGAATTA ATTGGAACAG AAGAAATAAT CAAGGAGGAG 1440
 GAGAGGGGAA AAGACATTTA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500
 AACCAATACA GGAAGGAGGA ACCCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560
 ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620
 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680
 ACAGAAAAAG ATATTTCCTT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740
 GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAAGT TCTTAGATC TCCACATATG 1800
 AACTTGTCCG GACTGCGACA ATCTTAAAT ACAGTTTCTA TAACAGAAAT TGAGGAGGAG 1860
 AGTTTATTGA CCAAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAG CTCCAGTCCC 1920
 GCAACTTCTG CTATCCCATC CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980
 GAAACCCCA AGACATAAAC ATATGATGTC CTTATACCA AATCTGCTAG AAATGCTTCC 2040
 GAGATTCAA CTTCATCAGC TTCAGAGAA TCACTAAAGG ATCCTTCTAT GGAGGGAAAT 2100
 GTGTGTTTCT CTAGCTCTAC AGACATAACA GCACAGCCG ATGTTGATC AGCGAGAGAG 2160
 AGCTTTCTCC CACTTAATTA CACTGAGATA CGTGTGATG AATCTGAGAA GACAACCAAG 2220
 TCCTTTTCTG CAGGCCCATG GATGTCACAG GGTCCCTCAG TTACAGATCT GGAATGCCA 2280
 CATATTCTA CTTTCCCTTA CTTCCCAACT GAGGTAAAC CACATGCTTT TACCCCATCC 2340
 TCCAGACAA AGGATTGTGT CTCCACGGTC AACGTGGTAT ACTGCAGAC AACCCACCG 2400
 GTATACAATG AGGCCAGTAA TAGTAGCCAT GAGTCTCGTA TTGTTCTAGC TGAGGGGTTG 2460

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

GAATCCGAGA AGAAGGCAGT TATACCCCTT GTGATCGTGT CAGCCCTGAC TTTTATCTGT 2520
CTAGTGGTTC TTGTTGGTAT TCTCATCTAC TGGAGGAAAT GCTTCCAGAC TGCACACTTT 2580
TACTTAGAGG ACAGTACATC CCTAGAGTT ATATCCACAC CTCCAACACC TATCTTTCCA 2640
ATTTAGATG ATGTGCGAGC AATTCCTAATA AAGCACTTTC CAAAGCATGT TGCAGATTTA 2700
CATGCAAGTA GTGGGTTTAC TGAAGAATTT GAGGAAGTGC AGAGCTGTAC TGTGTACTTA 2760
GGTATTACAG CAGACAGCTC CAACCAOCCA GACAACAAGC ACAAGAATCG ATACATAAAT 2820
ATCGTTGCACT ATGATCATAG CAGGGTTAAG CTAGCACAGC TTGCTGAAAA GGATGGCAAA 2880
CTGACTGATT ATATCAATGC CAATTATGTT GATGGCTACA ACAGACCAAA AGCTTATATT 2940
GCTGCCCAAG GCCCACTGAA ATCCACAGCT GAAGATTCTT GGAGAATGAT ATGGGAACAT 3000
AATGTGGAAG TTATTGTTCAT GATAACAAAC CTCCTGGAGA AAGGAAGGAG AAAATGTGAT 3060
CAGTACTGGC CTGCGGATGG GAGTGAGGAG TACGGGAAC TTTCTGGTCA TCAGAAGAGT 3120
GTGCAAGTGC TTGCTTATTA TACTGTGAGG AATTTTACT TAAGAAACAC AAAAATAAAA 3180
AAGGGCTCCC AGAAGGAAG ACCCACTGGA CGTGTGGTCA CACAGTATCA CTACACGCAG 3240
TGGCTGACA TGGGAGTACC AGAGTACTCC CTGCCAGTGC TGACCTTTGT GAGAAAGGCA 3300
GCCATGCGCA AGCGCCATGC AGTGGGGCCT GTTGTCTGCC ACTGCAGTGC TGGAGTTGGA 3360
AGAACAGGCA CATATATTGT GCTAGACAGT ATGTTGCAGC AGATTCAACA CGAAGGAAC 3420
GTCAACATAT TTGGCTTCTT AAAACACATC CGTTCACAAA GAAATATTAT GGTACAAACT 3480
GAGGAGCAAT ATGCTTCTAT TCATGATACA CTGGTTGAGG CCATACTTAG TAAAGAAACT 3540
GAGGTGCTGC ACAGTCAAT TCATGCCTAT GTTAATGCAC TCCTCATTC TGGACCAGCA 3600
GGCAAAACAA AGCTAGAGAA ACAATTCCAG CTCCTGAGCC AGTCAATAT ACAGCAGAGT 3660
GACTATTCTG CAGCCCTAAA GCAATGCAAC AGGGAAAGA ATCGAACTTC TTCTATCATC 3720
CCTGTGAAAA GATCAAGGAT TGGCATTTC TCCCTGAGTG GAGAAGGCAC AGACTACATC 3780
AATGCTCCT ATATCATGGG CTATTACCAG AGCAATGAAT TCATCATTAC CCAGCACCTC 3840
CTCCTTCATA CCATCAAGGA TTTCTGGAG ATGATATGG ACCATAATGC CCAACTGGTG 3900
GTTATGATT CTGATGGCCA AAACATGGCA GAAGATGAAT TTGTTTACTG GCCAAATAAA 3960
GATGAGCCTA TAAATTGTGA GAGCTTTAAG GTCACCTTA TGGCTGAAGA ACACAAATGT 4020
CTATCTAATG AGGAAAAATG TATAATTGAG GACTTTATCT TAGAAGCTAC ACAGGATGAT 4080
TATGACTTGC AAGTAGGACA CTTCAGTGT CCTAAATGGC CAAATCCAGA TAGCCCCATT 4140
AGTAAAACTT TTGAACCTAT AAGTGTATA AAAGAAGAAG CTGCCAATAG GGATGGGCTC 4200
ATGATTGTT ATGATGAGCA TGGAGGAGT ACGGCAGGAA CTTTCTGTGC TCTGACAACT 4260
CTTATGCACC AACTAGAAAA AGAAATTCCT GTGAGTGTTC ACCAGGTAGC CAAGATGATC 4320
AATCTAGAGG GGCAGGAGT CTTTGTGAC ATTGAGCAGT ATCAGTTTCT CTACAAAGTG 4380
ATCCTCAGCC TTGTGAGCAC AAGGCAGGAA GAGAATCCAT CCACCTCTCT GGACAGTAAT 4440
GGTGCAGCAT TGCCGTAGTG AAATATAGCT GAGAGCTTAG AGTCTTAGT TTAACACAGA 4500
AAGGGGTGGG GGGACTCACA TCTGAGCATT GTTTCTCTCT TCCTAAAAAT AGGCAGGAAA 4560
ATCAGTCTAG TTCTGTATTG TGTGATTTT CCATCACCTG ACAGTAACCT TCATGACATA 4620
GGATTCTGCC GCCAAATTTA TATCATTAA AATGTGTGCC TTTTGTCAAG ACTTGTAAAT 4680
TACTTATTAT GTTTGAACTA AAATGATTGA ATTTTACAGT ATTCTAAGA ATGGAATTGT 4740
GGTATTTTT TCTGTATTGA TTTTAACAGA AAATTTCAAT TTATAGAGT TAGGAATTCC 4800
AAACTACAGA AAATGTTTGT TTTTAGTGTC AAATTTTGT CTGTATTGT AGCAATTATC 4860
AGGTTTGCTA GAAATATAAC TTTTAATACA GTAGCCTGTA AATAAACAC TCTTCCATAT 4920
GATATTCAAC ATTTTCAAC TGCAGTATTC ACCTAAAGTA GAAATAATCT GTTACTTATT 4980
GTAATATCTG CCCTAGTGTG TCCATGGACC AAATTTATAT TTATAATGT AGATTTTTAT 5040
ATTTTACTAC TGAGTCAAGT TTTCTAGTTC TGTTAATTG TTTAGTTTAA TGACGTAGTT 5100
CATTAGCTGG TCTTACTCTA CCAGTTTCTT GACATTGTAT TGTGTTACCT AAGTCATTAA 5160
CTTTGTTTCA GCATGTAAAT TTAACTTTGT TGGAAAAATG AAATACCTTC ATTTTGAAG 5220
AAGTTTTTAT GAGAAATAAC CCTTACAAA CATTTGTTCA ATGGTTTTTA TCCAAGGAAT 5280
TGCAAAATAA AATATAAATA TTGCCATTAA AAAAAAATAA AAAAAAATAA 5340
AAA

Seq ID NO: 184 Protein sequence:
Protein Accession #: EOS sequence

55
60
65
70
75
80

1 11 21 31 41 51
MRILKRFLAC IQLLCVCRLD WANGYYRQR KLVEEIGWSY TGALEQKNWG KKYPTCNSPK 60
QSPINIDEDL TQVNVNKLKL KFGWDKTSI ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120
FKASKITHFW GKCNMSDDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEAB KKGKLRALS 180
ILEVGTREN LDFKALIDGV ESVSRFGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240
TDTVDWIVFK DTVSISESQL AVFCEVLTMQ QSGYVMLMDY LQNNFREQQY KPSRQVFSY 300
TGKEBIHEAV CSSEPNVQA DPENYTSLLV TWERPRVVD TMIEKFAVLY QQLDGEDQTK 360
HEFLTGDYKD LGAILNNLLP NMSYVLQIVA ICTNGLYKGY SDQLIVDMPT DNPELDLPPE 420
LIGTEEIIKE EEEGKDIEEG AIVNPGRDSA TNQIRKKEPQ ISTTHYNRI GTKYNEAKTN 480
RSPTRGSEFS KGDVVENTSL NSTSQPVTKL ATEKDILSLT QVTLELPPHT VEGTSASLND 540
GSKTVLRSPH MNLGSAESL NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600
ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDFSMEG NVWPPSSTDI 660
TAQPDVVGSR ESFLQNTYTE IRVDESEKTT KSPSAGPVMS QGSPVTDLEM PHYSTFAYFP 720
TEVTPHAFTP PSRQQLVST VNVVYSQTQ PVYNEASNSS HESRIGLAEG LESEKKAVIP 780
LVIVSALTFI CLVVLVGLI YWRKCFQTAH FYLEDSTSPR VISTPPTPIF PISDDVGAIP 840
IKHFPKHVAD LHASSGFTEE FEEVQSCVD LGITADSSNH PDKHKNRYI NIVAYDHSRV 900
KLAQLAEKDG KLTIDYINANY VDGYNRPKAY IAAQGLPKST AEDFWRMIWE HNEVIVMIT 960
NLVERGRKRC DQVWPADGSE EYGNFLVTQK SVQVLAYTV RNFTLRNTKI KKGSKGRPS 1020
GRVVTQHYHT QWPDMSVPEY SLPVLTFVRK AAYAKRHAVG PVVHCSAGV GRTGTIYILD 1080
SMLQQIQHEG TWNIQFGLKH IRSQNYLVQ TEBQVFIHD TLVEAILSKE TEVLDSHIHA 1140
YVNALLIPGP AGTKLEKQF QLLSQSNIQQ SDYSAALKQC NREKNRTSSI IPVERSRVGI 1200
SSLSGEGTDY INASYIMGY QSNFPIITQH PLLHTIKDFW RMIWDHNAQL VVMIPDGNM 1260
AEDEFVYVFN KDEPINCESF KVTLMABEHK CLSNEEKLI QDFILEATQD DVLVEVRHFQ 1320
CPKPNPDPSP ISKTPELISV IKEEAANRDG PMIVHDEHGG VTAGTFCALT TLMHOLEKEN 1380
SVDVYQVAKM INLMRPGVFA DIBQYQFLYK VILSLVSTRQ EENPSTSLDS NGAALPDGNI 1440
AESLESLSV

Seq ID NO: 185 DNA sequence
Nucleic Acid Accession #: EOS sequence
Coding sequence: 501-4514

85
1 11 21 31 41 51

	CACACATAGG	CACGACGAT	CTCACTTGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCCCTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
5	CGGCGAGGGG	CCGCGAGCCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCTC	CGCTTGCATT	180
	CAGCTCCTCT	GTGTTTGGCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAT	TGGGGAAAGA	300
	AATATCCAAT	ATGTAATAGC	CCAAAACAAT	CTCCTATCAA	TATTGATGAA	GATCTTACAC	360
	AAGTAAATGT	GAATCTTAAG	AAACTTAAAT	TTCAGGGTTG	GGATAAAACA	TCATTGGAAA	420
10	ACACATTCAAT	TCATAACACT	GGGAAAACAG	TGGAAATTA	TCTCACTAAT	GACTACCGTG	480
	TCAGCGGAGG	AGTTTCAGAA	ATGGTGTTTA	AAGCAAGCAA	GATAACTTTT	CAGTGGGGAA	540
	AATGCAATAT	GTCACTGAT	GGATCAGAGC	ATAGTTTAGA	AGGACAAAAA	TTTCCACTTG	600
	AGATGCAAT	CTACTGCTTT	GATGCGGACC	GATTTTCAAG	TTTTGAGGAA	GCAGTCAAG	660
	GAAAAGGGAA	GTTAAGAGCT	TTATCCATTT	TGTTTGAAGT	TGGGACAGAA	GAAAATTTGG	720
	ATTTCAAAAGC	GATTATTGAT	GGAGTCGAAA	GTGTTAGTGG	TTTTGGGAAG	CAGGCTGCTT	780
15	TAGATCCATT	CATACCTGTT	AACTTCTGCG	CAAACTCAAC	TGACAAGTAT	TACATTTACA	840
	ATGGCTCATT	GACATCTCCT	CCCTGCACAG	ACACAGTTGA	CTGGATTGTT	TTTAAAGATA	900
	CAGTTAGCAT	CTCTGAAAGC	CAGTTGCGTG	TTTTTTGTGA	AGTTCTTACA	ATGCAACAAT	960
	CTGGTTATGT	CATGCTGATG	GACTACTTAC	AAAACAATT	TGAGAGCAG	CAGTACAGT	1020
20	TCTCTAGACA	GGTGTTTTCC	TCATACACTG	GAAAGGAAGA	GATTCTAGAA	GCAGTTTGT	1080
	GTTCCAGAAC	AGAAAATGTT	CAGGCTGACC	CAGAGAAATTA	TACCAGCCTT	CTTGTTTACAT	1140
	GGGAAAGACC	TCCAGTCGTT	TATGATACCA	TGATTGAGAA	GTTTGCAGTT	TTGTACCAGC	1200
	AGTTGGATGG	AGAGGACCAA	ACCAAGCATG	AATTTTGTAC	AGATGGCTAT	CAAGACTTGG	1260
	GTGCTATTCT	CAATAATTTG	CTACCCAATA	TGAGTTATGT	TCTTCAGATA	GTAGCCATAT	1320
	GCCTAATGG	CTTATATGGA	AAATACAGCG	ACCAACTGAT	TGTGACATG	CCTACTGATA	1380
25	ATCTGAAGC	TGATCTTTTC	CCTGAATTA	TTGGAATCGA	AGAAAATAAT	AAGGAGGAGG	1440
	AAGAGGGAAA	AGACATTGAA	GAGGGCGCTA	TTGTGAATCC	TGGTAGAGAC	AGTGCTACAA	1500
	ACCAAAATCAG	GAAAAAGGAA	CCCCAGATT	CTACCAACA	ACACTACAAT	CGCATAGGGA	1560
	CGAAATACAA	TGAAGCCAG	ACTAACCGAT	CCCCAACAG	AGGAAGTGAA	TTCTCTGGAA	1620
	AGGGTGATGT	TCCCAATACA	TCTTTAAAT	CCACTTCCCA	ACCACTGCT	AAATTAGCCA	1680
30	CAGAAAAGCA	TATTTCTTGG	ACTTCTCAGA	CTGTGACTGA	ACTGCCACCT	CACACTGTGG	1740
	AAGGTACTTC	AGCCTCTTTA	AATGATGGCT	CTAAACTGT	TCTTAGATCT	CCACATATGA	1800
	ACTTGTGCGG	GACTGCAGAA	TCCTTAAATA	CAGTTTCTAT	AACAGAAAT	GAGGAGGAGA	1860
	GTPTATTGAC	CAGTTTCAAG	CTTGATACTG	GAGCTGAAGA	TTCTTCAGGG	TCCAGTCCCG	1920
35	CAACTTCTGC	TATCCCATTC	ATCTCTGAGA	ACATATCCCA	AGGGTATATA	TTTTCTCTCG	1980
	AAAACCCAGA	GACATAACCA	TATGATGTCC	TTATACCAGA	ATCTGCTAGA	AATGCTTCCG	2040
	AAGATTCAAC	TTCATCAGGT	TCAGAGAAGT	CACATAACAG	CACAGCCCGA	TGTTGGATCA	2100
	TGTGGTTTCC	TAGCTCTACA	GACTAGATAC	GTGTTGATGA	ATCTGAGAAG	ACAACCAAGT	2160
	GCCTTCTCCA	GACTAATTTG	ACTGAGATAC	GTGCTGATGA	ATCTGAGAAG	ACAACCAAGT	2220
40	CCCTTCTGAC	AGGCCAGTGG	ATGTCACAGG	GTCCCTCAGT	TACAGATCTG	GAAATGCCAC	2280
	ATTATCTTAC	CTTTGCTTAC	TTCCCAACTG	AGGTAACACC	TCATGCTTTT	ACCCCATCCT	2340
	CCAGACAACA	GGATTGGTTC	TCCACGGTCA	ACGTGGTATA	CTGCGAGACA	ACCCCAACCG	2400
	TATACAATGA	GGCCAGTAAT	AGTAGCCATG	AGTCTGATAT	TGGTCTAGCT	GAGGGGTTGG	2460
	AATCCGAGAA	GAAGGCAGTT	ATACCCCTTG	TGATCTGTGC	AGCCCTGACT	TTTATCTGTC	2520
45	TAGTGGTTCT	TGTGGGTATT	CTCATCTACT	GGAGGAAATG	CTTCCAGACT	GCACACTTTT	2580
	ACTTAGAGCA	CAGTACATCC	CCTAGAGTTA	TATCCACACC	TCCACACCTT	ATCTTTCCAA	2640
	TTTCAGATGA	TGTGCGAGCA	ATTCCAATAA	AGCACTTTCC	AAAGCATGTT	GCAGATTTAC	2700
	ATGCAAGTAG	TGGGTTTACT	GAAGAATTTG	AGACACTGAA	AGAGTTTAC	CAGGAAGTGC	2760
	AGAGCTGTAC	TGTTGACTTA	GGTATTACAG	CAGACAGCTC	CAACCAACCA	GACAAACAAG	2820
50	ACAAGAATCG	ATACATAAAT	ATCGTTGCC	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	2880
	TTGCTGAAGA	GGATGGCAAA	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	2940
	ACAGACCAAA	AGCTTATATT	GCTGCCCAAG	CCCCACTGAA	ATCCACAGCT	GAAGATTTC	3000
	GGAGAATGAT	ATGGGAACAT	AATGTGGAAG	TTATTGTGAT	GATAACAAC	CTCGTGGAGA	3060
	AAGGAAGGAG	AAAATGTGAT	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAAC	3120
55	TTCTGGTCA	TCAGAGAGGT	GTGCAAGTGC	TGCGCTATTA	TACTGTGAGG	AAITTTACTC	3180
	TAAGAAACAC	AAAAATAAAA	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGA	CGTGTGGTCA	3240
	CACAGTATCA	CTACAGCGCA	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	3300
	TGACCTTTGT	GAGAAAGGCA	GCCTATGCCA	AGCGCCATGC	AGTGGGCTCT	GTGTTGCTGC	3360
	ACTGCACTGC	TGGAGTTGGA	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	3420
60	AGATTCAACA	CGAAGGAAC	GTCAACATAT	TGCGCTTCTT	AAAACACAT	CGTTCACAAA	3480
	GAAATATTAT	GGTACAACT	GAGGAGCAAT	ATGCTTCTAT	TCATGATACA	CTGGTTGAGG	3540
	CCATACCTAG	TAAAGAACT	GAGGTGCTGG	ACAGTCAAT	TCATGCTAT	GTTAATGCAC	3600
	TCCTCATTC	TGGACCAAG	GGCAAAACAA	AGCTAGAGAA	ACAATTCAG	CTCCTGAGCC	3660
	AGTCAAAAT	ACAGCAGAGT	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAAGAA	3720
65	ATCGAACTTC	TTCTATCATC	CCTGTGGAAA	GATCAAGGGT	TGGCAATTC	TCCCTGAGTG	3780
	GAGAAGGCAC	AGACTACATC	AATGCCTCCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	3840
	TCATCATTAC	CCAGCACCTC	CTCCTTACAT	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	3900
	ACCATAATGC	CCAACCTGGT	GTTATGATTC	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	3960
	TGTTTACTG	GCCAAATAAA	GATGAGCCTA	TAAATTTGTA	GAGCTTTAAG	GTCACCTTAA	4020
70	TGGCTGAAGA	ACACAAATGT	CTATCTAATG	AGGAAAAACT	TATAATTCAG	GACTTTATCT	4080
	TAGAAGCTAC	ACAGGATGAT	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	4140
	CAATCCAGA	TAGCCCCATT	AGTAAACTT	TTGAACCTAT	AAGTGTATA	AAAGAAGGAG	4200
	CTGCCAATAG	GGATGGGCTT	ATGATTGTTT	ATGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	4260
	CTTCTGTGC	CTTGACAACC	CTTATGCACC	AACATAGAAA	AGAAAATTC	GTGGATGTTT	4320
75	ACCAAGGTAG	CAAGATGATC	AATCTGATGA	GGCCAGGAGT	CTTGTGCTGAC	ATTGAGCAGT	4380
	ATCAGTTTCT	CTACAAAGTG	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	4440
	CCACCTCTCT	GGACAGTAAT	GGTGCGAGAT	TGCCCTGATG	AAATATAGCT	GAGAGCTTAG	4500
	AGTCTTTAGT	TAAACACAGA	AAGGGGTGGG	GGGACTCACA	TCTGAGCAAT	GTTTTCTCT	4560
	TCCTAAAATT	AGGCAGGAAA	ATCAGTCTAG	TCTGTTATC	TGTTGATTTC	CCATCACCTG	4620
80	ACAGTAACTT	TCATGACATA	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	4680
	TTTTTCAAG	ACTTGTAAAT	TACTTATTAT	GTTTGAACCTA	AAATGATTGA	ATTTTACAGT	4740
	ATTTCTAAGA	ATGGAATTTG	GGTATTTTTT	TCTGTTATGA	TTTTAACAGA	AAATTTCAAT	4800
	TTATAGAGGT	TAGGAATTTCC	AAACTACAGA	AAATGTTTGT	TTTTAGTGT	AAATTTTAG	4860
	CTGTATTGT	AGCAATTATC	AGGTTTGCTA	GAAATATAAC	TTTTAATACA	GTAGCCTGTA	4920
	AAATAAACAC	TCTTCCATAT	GATATTCAAC	ATTTTACAAC	TGCAGTATTC	ACCTAAAGTA	4980
85	GAAATAACT	GTTACTTATT	GTAATACTG	CCCTAGTGT	TCCATGGACC	AAATTTATAT	5040
	TTATAATTGT	AGATTTTAT	ATTTTACTAC	TGAGTCAAGT	TTTCTAGTTC	TGTGTAATG	5100
	TTTAGTTTAA	TGACGTAGTT	CATTAGCTGG	TCTTACTCTA	CCAGTTTCT	GACATTGTAT	5160

TGTTTACCTT AAGTCATTAA CTTTGTITCA GCATGTAATT TTAACITTTG TGGAAAAATAG 5220
 AAATACCTTC ATTITGAAAG AAGTTTITAT GAGAATAACA CCTTACCATA CATTTGTTCAA 5280
 ATGGTTTTTA TCCAAGGAAT TGCAAAAAATA AATATAAATA TTGCCATTAA AAAAAAATAA 5340
 AAAAAAATAA AAAAAAATAA AAA

Seq ID NO: 186 Protein sequence:
 Protein Accession #: EOS sequence

1 11 21 31 41 51
 MVFKASKITP HWGKCNMSSD GSEHSLEGQK FPLEMQIYCP DADRFSSFEE AVKKGKGLRA 60
 LSILFEVGTB ENLDFKAIID GVESVSRFGK QAALDPFILL NLLFNSTDKY YIYNGSLTSP 120
 PCTDVTVDWIV FKDTVSISES QLAVFCEVLT MQQSGYVMLM DYLNQNFREQ QYKFSRQVFS 180
 SYTGKEEIH EAVCSSEFENV QADPENYTSI LVTWERPRV VDTMIEKFAV LYQQLDGEDQ 240
 TKHEFLT DGY QDLGAILNNE LPNMSYVLQI VAICTNGLYG KYSDQLIVDM PTDNPELDF 300
 PELIGTEBII KEEBEGKDI EGAIVNPGRD SATNQIRKKE PQISTTTHYN RIGTKYNEAK 360
 TNRSPTRGSE PSGKGDVPT SLNSTSQPVT KLATEKDIDL TSQVTLEPP HTVEGTSASL 420
 NDGSKTVLRV PBMNLSGTAE SLNVT SITEY EESLLTSFK LDTGAEDSSG SSPATSAIFP 480
 ISENIQGYI FSSENPETIT YDVLIPESAR NASEDSTSSG SEESLKDPSM EGNVWFPSST 540
 DITAPQDVGS GRESFLOQNY TEIRVDESEK TTKSFSAGPV MSQGPSVTDL EMPHYSTFAY 600
 FPTEVTPHAF TPSSRQQLDV STVNVSYSQT TQPVYNEASN SSHERIGLA EGLESEKKAV 660
 IPLIVYSALT FICLVVLVGI LIYMRKCFQT AHFYLEDSTS PRVISTPPT IPFISDDVGA 720
 IPIKFFPHV ADLHASSGFT EEFETLKEFY QEVQSCITVDL GITADSSNHP DNKHKNRYN 780
 IVAYDHSRVK LAQLAEKDGK LDTYINANYV DGYNRPKAYI AAQGPLKSTA EDFWRMIWEH 840
 NVEIVMITN LVEKGRKCD QYWPADGSEE YGNFLVTQKS VQVLAYYTVR NFTLRNTKIK 900
 KGSQKGRPSG RVVTVYHYTQ WPDMDGPEYS LPVLTFVRKA AYAKRHAVGP VVHCSAGVG 960
 RTGTIYLDL FICLVVLVGI VNIFFGLKHI RSQRNVLVQT EEQVFIHDT LVEAILSKET 1020
 EVLDSHIRAY VNALLIPGFA GKTKLEKQFO LLSQSNIQQS DYSAAKQCN REKNRTSSII 1080
 PVERSRVGIS SLGSECTDYI NASYIMGYVQ SNEFIITQHP LLHTIKDFWR MIWDHNAQLV 1140
 VMIPDQDMA EDEFVYWPKN DEPINCESFK VTLMAEEHK LSNEBKLIQ DFLEATQDD 1200
 YVLEVRHQC PKWPNPDSPI SKTFELISVI KEAANRNDGP MIVDEHGGV TAGTFCALTT 1260
 LMHQLKENS VDVYQVAKMI NLMRPGVFAD IEQYQFLYKV ILSLVSTRQE ENPSTSLDSN 1320
 GAALPDGNIA ESLESIV

Seq ID NO: 187 DNA sequence
 Nucleic Acid Accession #: EOS sequence
 Coding sequence: 148-4632

1 11 21 31 41 51
 CACACATACG CACGACACGAT CTCACITCGA TCTATACACT GGAGGATTAA AACAAACAAA 60
 CAAAAAACA ATTTCCTTCG CTCCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120
 CGGCGAGGGG CGCAGACCGC TCTGGAATG CGAATCCTAA AACGTTTCCT CGCTTGCAAT 180
 CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
 CTTGTGAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAAA TTGGGGAAAG 300
 AAATATCCAA CATGTAATAG CCAAAACAAA TCTCCTATCA ATATTGATGA AGATCTTACA 360
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAACC ATCATGGGAA 420
 AACACATCA TTCAATAACAC TGGGAAAAA GTGGAAATTA ATCTCACTAA TGACTACCGT 480
 GTCACCGGAG GAGTTTTCAG AATGGTGTTT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540
 AAATGCAATA TGTACTCTGA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCACACT 600
 GAGATGCAAA TCTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660
 GGAAAAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAAG TTGGGACAGA AGAAAATTGG 720
 GATTTCAAGC CGATTATTGA TGGAGTCGAA AGTGTAGTGC GTTTTGGGAA GCAGGCTGCT 780
 TTGATCATCT CATACTGTTT GAACCTCTCG CCAAACTCAA CTGACAAGTA TTACATTTAC 840
 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTGG ACTGGATTGT TTTTAAAGAT 900
 ACAGTTAGCA TCTCTGAAGC CAGTTGGCTT GTTTTGTGAG AAGTTCTTAC AATGCAACAA 960
 TCTGTTATG TCACTGCTAT GGACTACTTA CAAAACAATT TCGAGAGCA ACAGTACAAAG 1020
 TTCTCTAGAC AGGTGTTTTT CTCATACACT GGAAGGAAG AGATTCTAGA AGCAGTTTGT 1080
 AGTTCAGAAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCTT TCTGTGTTCA 1140
 TGGGAAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGAGT TTTGTACCAAG 1200
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260
 GGTGCTATCT TCAATAATTG GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320
 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTCGACAT GCCTACTGAT 1380
 AATCTCGAAC TTGATCTTTT CCTGGAATTA AITGGAACCT AAGAATAAT CAAGGAGGAG 1440
 GAAGAGGGAA AAGACATTGA AAGAGGCGCT AITGTGAATC CTGGTAGAGA CAGTGTCTACA 1500
 AACCAATCA GGAAGAAAGGA ACCCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560
 ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCCACAAA GAGGAAGTGA ATTCTCTGGA 1620
 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680
 ACGAAAAAAG ATATTCTCTT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740
 GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAACCTG TTCTTAGATC TCACATATG 1800
 AACTTGTGCG GACTTGCAGA ATCCTTAAAT ACAGTTTCTA TAACAGATA TGAGGAGGAG 1860
 AGTTTATTGA CCAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920
 GCAACTTCTG CATCTCCATT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980
 GAAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC 2040
 GAAGATTCAA CTTTATCAGG TTCAGAAGAA TCACTAAAGG ATCCTTCTAT GGAGGGAAT 2100
 GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCG ATGTTGGATC AGGCAGAGAG 2160
 AGCTTTCTCC CACTGAATTA CACTGAGATA CGTGTGATG AATCTGAGAA GACAACCAAG 2220
 TCCTTTCTG CAGGCCAGT GATGTACAGG GGTCCCTCAG TTACAGATCT GGAATGCCA 2280
 CATTATTCTA CCTTTCCTTA CTTCCCAACT GAGGTAACAC CTCATGCTTT TACCCCATCC 2340
 TCCAGACAAC AGGATTGGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCAACCG 2400
 GTATACAAAT AGGCCAGTAA TAGTAGCCAT GAGTCTCGTA TTGGTCTAGC TGAGGGGTTG 2460
 GAATCCGAGA AGAAGGCAGT TATACCCCTT GTGATCGTGT CAGCCCTGAC TTTTATCTGT 2520
 CTAGTGGTTC TTGTGGGTAT TCTATCTAC TGGAGGAAAT GCTTCCAGAC TGCACTTTT 2580
 TACTTAGAGG ACAGTACATC CCTTAGAGTT ATATCCACAC CTCACACCC TATCTTTCCA 2640
 ATTTGAGATG ATGTCGAGC AATTCCAATA AAGCACTTTC CAAAGCATGT TGCAGATTTA 2700
 CATGCAAGTA GTGGGTTTAC TGAAGAATTT GAGACACTGA AAGAGTTTTA CCAGGAAGTG 2760
 CAGAGCTGTA CTGTTGACTT AGGTATTACA GCAGACAGCT CCAACCACCC AGACAACAG 2820

5
10
15
20
25
30
35
40
45

```

CACAGAATC, GATACATAAA TATCGTGGC TATGATCAT GCAGGGTTAA GCTAGCACAG 2880
CTTGTCTGAAA AGGATGGCAA ACTGACTGAT TATATCAATG CCAATTATGT TGATGGGTAC 2940
AACAGACCAA AAGCTTATAT TGCTGCCCAA GGCCCACTGA ATCCACAGC TGAAGATTTC 3000
TGGAGAAATGA TATGGGAACA TAATTTGGAA GTTATTGTCA TGATAACAAA CCTCGTGGAG 3060
AAAGGAAGGA GAAATTTGTA TCAGTACTGG CCTGCGGATG GGAGTGAGGA GTACGGGAAC 3120
TTTCTGGTCA CTCAGAAAGG TGTGCAAGTG CTTGCCATT ATACTGTGAG GAATTTTACT 3180
CTAAGAAACA CAAAATAAAA AAAGGGCTCC CAGAAAGGAA GACCCAGTGG ACCTGTGGTC 3240
ACACAGTATC ACTACACGCA GTGGCCTGAC ATGGGAGTAC CAGAGTACTC CCTGCCAGTG 3300
CTGACCTTTG TGAGAAAGGC AGCCTATGCC AAGCGGCATG CAGTGGGGCC TGTGTCTGTC 3360
CACTGCAGTG CTGGAGTTGG AAGAACAGGC ACATATATTG TGCTAGACAG TATGTTGCAG 3420
CAGATTCAAC AGAAGGAAAC TGTCAACATA TTTGGCTTCT TAAACACAT CCGTTCACAA 3480
AGAAATTATT TGGTACAAAC TGAGGAGCAA TATGTCTTCA TTCAATGATC ACTGGTTGAG 3540
GCCATACTTA GTAAAGAAAC TGAGGTGCTG GACAGTCATA TTCAATGCTA TGTATATGCA 3600
CTCTCTATTC CTGGACCAGC AGGCAAAACA AAGCTAGAGA AACAATTCCA GGGTCTCACT 3660
CTGTCAACCA GGCTGGAGTG CAGAGGCAAC ATCTGGGCTC ACTGCAACCT TCCTCTCCCT 3720
GGCTTAATCT ATCTCTCTAC CTGAGCCTCC CGAGTGGGCT GGAATATACT CCTGAGCCAG 3780
TGCAATATAT AGCAAGTGA CTATTCTGCA GGCCTAAAGC AATGCAACAG GGAAGAAGAT 3840
CGAATTCTCT CTATCATCCC TGTGGAAGA TCAAGGGTTG GCATTTCATC CCTGAGTGGG 3900
GAAGGCACAG ACTACATCAA TGCTCTCTAT ATCATGGGCT ATTACCAGAG CAATGAATTC 3960
ATCATTATCC AGCACCCTCT CCTTCATACC ATCAAGGATT TCTGGAGGAT GATATGGGAC 4020
CATATGTGCC AACTGGTGGT TATGATTCTC GATGGCCAAA ACATGGCAGA AGATGAATTT 4080
GTTTACTGGC CAAATAAAGA TGAGCCTATA AATTGTGAGA GCTTTAAGGT CACTCTTATG 4140
GCTGAGAAC ACAAATGTCT ATCTAATGAG GAAAACTTA TAATTGAGG CTTTATCTTA 4200
GAAGCTACAC AGGATGATTA TGTACTTGA GTGAGGCACT TTCAGTGTCC TAAATGGCCA 4260
AATCCAGATA GCCCATTAG TAAACTTTT GAACCTATAA GTGTTATAAA AGAAGAAGCT 4320
GCCAATAGGG ATGGCCCTAT GATTGTTCAT GATGAGCATG GAGGAGTAC GGCAGGAAGT 4380
TTCTGTGCTC TGACAACCTC TATGCCCAA CTAGAAAAAG AAAATTCCGT GGATGTTTAC 4440
CAGGTAGCCA AGATGATCAA TCTGATGAGC CCAGGAGTCT TTGCTGACAT TGAGCAGTAT 4500
CAGTTTCTCT ACAGAGTATG CCTCAGCCTT GTGGGCACAA GGCAGGAAGA GAATCCATCC 4560
ACCTCTCTGG AGCAATATGG TGCAGCATTG CCTGATGGAA ATATAGCTGA GAGCTTAGAG 4620
TCTTTAGTTT AACACAGAAA GGGGTGGGGG GACTCACATC TGAGCATTGT TTCTCTCTTC 4680
CTAAAATTAG CGAGAGAAAT CAGTCTAGTT CTGTTATCTG TTGATTTCCT ATCACTGAC 4740
AGTAACCTTC ATGACATAGG ATTCTGCCGC CAAATTTATA TCATTACAA TGTTGCTCTT 4800
TTTGCAAGAC TTGTAATTTA CTTATTATGT TTGAACATA ATGATTGAAT TTTACAGTAT 4860
TTCTAAGATG GGAATTGGG TATTTTCTT TGTATTGATT TTAACAGAAA ATTTCAATTT 4920
ATAGAGGTTA GGAATTCCAA ACTACAGAAA ATGTTTGTTC TTAGTGTCAA ATTTTAGCT 4980
GTATTTGTAG CAATTATCAG GTTTGCTAGA AATATAACTT TTAATACAGT AGCCTGTAAA 5040
TAAACACTC TCCTATATGA TATTCAACAT TTTCAACTG CAGTATTCAC CTAAGTAGA 5100
AATAATCTGT TACTTATGT AATACTGCC CTAGTGTCTC CATGGACCAA ATTTATATTT 5160
ATAATTGTAG ATTTTATAT TTTACTACTG AGTCAAGTTT TCTAGTTCTG TGTAAATGTT 5220
TAGTTTAATG AGTATGTTCA TTAGCTGGTC TTACTCTACC AGTTTCTGA CATTTGATTT 5280
TGTTACTTAA GTCAATTAAC TTGTTTACG ATGTAATTTT AACTTTTGTG GAAATAGAAA 5340
ATACCTTCAT TTTGAAGAAA GTTTTATGTA GAATAACACC TTACCAACAA TTGTTCAAT 5400
GGTTTATATC CAGGAATTTG CAAAATAAAA TATAATATT GCCATTAAAA AAAAAAATA 5460
AAAAAATAA AAAAAAATA A

```

Seq ID NO: 188 Protein sequence:
Protein Accession #: E09 sequence

50
55
60
65
70
75

```

1 11 21 31 41 51
MRILKRFLAC IQLLCVCRLD WANGYYRQOR KLVEEIGWSY TGALEQKNWG KKYPTCNSPK 60
QSPINIDEDL TQVNVNKKLL KFGWDKTSL ENTFIHNIGK TVEINLTNDY RVSGGVSEMV 120
FKASKITFHW GKCNMSSDGS EHSLEGKFP LEMQIYCPDA DRFSFEEAV KKGKLRALS 180
ILFEVGTBEN LDFKAIIDGV ESVSFRGQA ALDPFILLNL LPNSTDYKYI YNGSLTSPPC 240
TDTVDWIVFK DTVSISBSQL AVFCEVLTMQ QSGVYMLMDY LQNNFREQQY KFSRQVFSY 300
TKKEBIHEAV CSSEPNVAVL DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360
HEFLTDGYQD LGAILNLLP NMSYVLQIVA ICTNGLYCKY SDQLIVDMPT DNPDLDFPE 420
LIGTEBIIKE ESEKDIIEG AIVNPRGDSA TNQIRKKEPQ ISTTTHYNNR GTRYNEAKTN 480
RSPTRGSSFS RSGDVPNTSL NSTSQPVTKL ATEKDISLTS QTVTELPPTT VEGTSASLND 540
GSKTVLRSFH MNLGSAESL NTVSITYEYE ESLLTSPFKL TGAEDSSGSS PATSAIPPIS 600
ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSGSGE ESLKDPSEMEG NVWFPSSTDI 660
TAQPDVSGSR ESFLQNTYNE IRVDESEKIT KSFSAGPVMS QGFSVTDLEM PHYSTFAYFP 720
TEVTPHAPT P SSRQQLVST VNVVYSQTQ PVYNASNS HESRIGLAE LESEKAVIP 780
LVIVSALTFI CLVVLVGLI YWRKCFQTAH FYLEDSTSPR VISTPPTPIF PISDDVGAIP 840
IKHFFKHVLA LHASGFTFE FETLKEFYQE VQSCVVDLGI TADSSNHPDN KHKRYINIV 900
AYDHSRVLAD QLAEDKGLT DYINANYVDG YNRPKAYIAA QGLKSTAE FWRMIWEHNV 960
EVIVMITNLV EKGRKCKDQY WPADGSEZYG NFLVTQKSVQ VLAYYTVRNF TLRNTKIKG 1020
SQKGRPSGRV VTQYHYTQWP DMGVPEYSLP VLFVFRKAA AKRHAVGPV VHCAGVGR 1080
GTIYIVLDSML QQIQHEGTVN IFGLKHRS QRNYLVQTEE QYVFIHDTLV EALLSKETE 1140
LDSHIBAVN ALLIPGPAK TKLEKQFQGL TSLPRLCRG TISAHCNPL PGLTDPPTSA 1200
SRVAGTILLS QSNIQSDYS AALKQCNREK NRTSSIIPE RSRVGISLS GEGTDYINAS 1260
YIMGYQSNF FIITQPHLLH TIKDFWRMW DHNAQLVMI PDGQNAEDE FVYWPKNDEP 1320
INCSEFKVTL MAEHLKCLN EEKLIQDFI LEATQDDYVL EVRHFQCPKW PNPDPISKT 1380
FELISVKEE AANRDGPNIV HDEHGGVTAG TPCALTLMH QLEKENSVDV YQVAKMINLM 1440
RPGVFADIEQ YQFLYKVLIS LVGTQRENP STSLDSNGAA LPDGNIAESL ESLV

```

Seq ID NO: 189 DNA sequence
Nucleic Acid Accession #: NM_002820
Coding sequence: 304..831

85

```

1 11 21 31 41 51
CGGTTGCA AGAAGCTGA CTTCAGAGG GGAACCTTC TTCTTTTAGG AGGCGGTTAG 60
CCCTGTTCCA CGAACCCAGG AGAAGCTGCT GCCAGATTAA TTAGACATTG CTATGGGAGA 120
CGTGAACAA CACTACTTAT CATTGATGCA TATATAAAC CATTTTATTG TCCTATTAT 180

```

5
 10
 15
 20
 25

```

TTCAGAGGAA GCGCCTCTGA TTGTTTCTT TTTTCCCTT TTGCTCTTC TGGCTGTGTG 240
GTTTGGAGAA AGCACAGTTG GAGTAGCCGG TTGCTAAATA AGTCCCGAGC GCGAGCGGAG 300
ACGATGTCAGC GGAGACTGGT TCAGCAGTGG AGCGTCGCGG TGTTCTGTCT GAGCTACGCG 360
GTGCCCCCTC GCGGGCGCTC GGTGGAGGGT CTCAGCCGCC GCCTCAAAAG AGCTGTGTCT 420
GAACATCAGC TCCTCCATGA CAAGGGGAAG TCCATCCAAG ATTTACGGCG ACCGATTCTTC 480
CTTCACCATC TGATCGCAGA AATCCACACA GCTGAAATCA GAGCTACCTC GGAGGTGTCC 540
CTTAACCTCA AGCCCTCTCC CAACACAAAG AACCACCCCG TCCGATTGGT GTCTGATGAT 600
GAGGGCAGAT ACCTAACTCA GGAACTAAC AAGGTGGAGA CGTACAAGA GCAGCCGCTC 660
AAGACACCTG GGAAGAAAAA GAAAGGCAAG CCCGGGAAAC GCAAGGAGCA GGAAGAGAAA 720
AAACGGCGAA CTCGCTCTGC CTGGTTAGAC TCTGGAGTGA CTGGGAGTGG GCTAGAAGGG 780
GACCACCTGT CTGACACCTC CACACGTCG CTGGAGCTCG ATTCACGGTA ACAGGCTTCT 840
CTGGCCCGTA GCCTCAGCGG GGTGCTCTCA GCTGGGTTT GGAGCCTCCC TTCTGCCTTG 900
GCTTGGACAA ACCTAGAATT TTCTCCCTT ATGTATCTCT ATCGATTGTG TAGCAATTGA 960
CAGAGAATAA CTCAGAATAT TGTCTGCTT AAAGCAGTAC CCCCTTACCA CACACACCCC 1020
TGTCCTCCAG CACCATAGAG AGGCGCTAGA GCCATTCTCT CTTTCTCCAC GGTCAACCCA 1080
CATCAATCTC TTACCATCTC ACCAAATAAT TTCATATCCA AGCTTCAGAA GCTAGTGACC 1140
ATCTTCATAA TTGCTGGAG AAGTGTATT CTCCCTTA CTCTCACACC TGGGCAAACT 1200
TTCTTCAGTG TTTTTCATT CTACGTTCT TTCACTCAA GGGAGAATAT AGAAGCATT 1260
GATATTATCT ACAACACTG CAGAACAGCA TCATGTCTA AACGATTCTG AGCCATTCTC 1320
ACTTTTATT TAATTAATTA TATTAAATTA AATCTCAAAT TTATTTTAA GTAAAGAACT 1380
TAAATATATG TTTAAACACA TGCCCTAAAT TTGTTTAAAT AAATTTAACT CTGGTTTCTA 1440
CCAGCTCATA CAAATAAAT GGTTCCTGAA AATGTTAAG TATTAACCTA CAAGGATATA 1500
GGTTTTTCTC ATGTATCTTT TTGTTTATTG GCAAGATGAA ATAATTTTTC TAGGGTAATG 1560
CCGTAGGAAA AATAAAACT CACATTTAAA AAAAA
  
```

Seq ID NO: 190 Protein sequence:
Protein Accession #: NP_002811

30
35

```

1 11 21 31 41 51
MQRRLLVQWS VAVFLLSYAV PSCGRSVEGL SRRLKRAVSE HQLLHDKGKS IQDLRRRFFL 60
HHLIASIHTA EIRATSEVSP NSKPSNPTRN HPVRFSGSDE GRYLTQETNK VETYKEQPLK 120
TPGKKKGKGP GKKEQEKKK RRTSAWLDS GVTGSGLEGD HLDSTSTTSL ELDSR
  
```

Seq ID NO: 191 DNA sequence
Nucleic Acid Accession #: XM_059328
Coding sequence: 52..1023

40
45
50
55
60
65

```

1 11 21 31 41 51
GGGCTGTCCG GCCCACTCCC CTGGGAGCGC GAGCGGTGGA CCCAGGCGGC CATGTCCCGC 60
CCTCGCATGC GCTCTGGTGT CACCGCGGAC GACTTTGGTT ACTGCCCGCG ACGCGATGAG 120
GGTATCGTGG AGGCTTTCTT GGCCGGGGCT GTGACCAGCG TGTCCTGTCT GGTCAACGGT 180
GCGGCCACGG AGAGCGCGGC GGAGCTGGCC CGCAGGCACA GCATCCCCAC GGGCCTCCAC 240
GCCAACTGTG CCGAGGGCGG CCGGTGGGT CCGGCCCGCC GTGGCGCCTC ATCGCTGCTC 300
GGCCCGGAAG GCTTCTCTCT TGGCAGATG GGATTCGGG AGGCGGTGCG GGCCGAGAC 360
GTGGATTTCG CTGAGGTGCG GGAGGAGCTC GAGGCCAAC TAAGCTGCTT CCGGAGCTG 420
CTGGGCAGGG CCCCCACGCA CGCGACGGG CACCAGCAG TGCACTGTCT CCGAGCGCTG 480
TGCCAGGTGT TGCCGAGGCG GCTGACGGCC TATGGGTGTC GCTTTACGCG ACTGCCGCTG 540
GAGCGGGTGT TGGGTGGCTG CACTTGGCTG GAGGCCCGCG CGCTGCTCTT CGCTGCGCC 600
GTGGAGCGCG AGCCCGGGCG CGCGTGGGC CCCTTCTCCC GCCACGGCT GCGGTGGACA 660
GAGCCCTTGG TGGGCTGAG CACTTGGCGC GGGCACATGT CCGCTCACCG CGTGTCCGG 720
GCCCTGGCGG GGTCTCTGGA AGGTACCTTA GCGGGCCACA CCTGACAGC GAGCTGATG 780
GCCGACCCCG GCTACCCAG TGTGCTTCCC ACCGCGGGCT GCGGTGAAGG CCGCGAGCT 840
TTCTCTTGT CTGGGAGGCG GCTGATGAG CTGCGCTGCC TCACCGCGCC CAGCTGCGG 900
GCCAGCTTG CCCAGGATGG CGTGACGCTT TGCGCCCTCG ACGACCTGGA CTCCAAGAGG 960
CCAGGGGAGG AGGTCCCCCT TGAGCCCACT CTGGAACCCCT TCCTGGAACC CTCCTACTC 1020
TGACCCCTTA CAGACAACCA AGCACTAATC CCCTTAGTAC CAAGAAAGGG GAGCCAGGAT 1080
TTAGTCTCTG CCGAGCCAG AGCTGGGACC TGGAGCAAGA TCTGTTGACT TCCTCTGGTA 1140
GGACACTGCC ACCTCTGGGC TCAGGTCTCT ATGCTTCCAA ATGGCATCTA GAGTTTGAGC 1200
AGCCTTCTTG GCTGACGGCA GGCCTAGCCT GTGGCAGCGG GCTAGGGCCC GCAGAGCATT 1260
TGTGCGCCCT CCATGTTGCA ATGCAACAC CTTACCACT GGGCAGTGG GGAGAGATGG 1320
CTATATTAAT AAAATAACGT GTGTCTTTC
  
```

Seq ID NO: 192 Protein sequence:
Protein Accession #: XP_059328

70
75

```

1 11 21 31 41 51
MSRPRMLVV TADDFGYCPR RDEGIVEAFL AGAVTSVSLI VNGAATESAA ELARRHSIPT 60
GLHANLSEGR PVGPARRGAS SLLGPEGFFL GKMGFREAVA AGDVLDPQVR EELEAQLSCF 120
RELLGRAPTH ADGHQHVLV PGVCQVPAEA LQAYGVRPTR LPLERGVGGC TWLEAPARAF 180
ACAVERDARA AVGPFSRHGL RWTDAFVGLS TCGRHMSAHR VSGALARVLE GTLAGHTLTA 240
ELMAHPGYPS VPPTGGCGEG PDAFSCSWER LHELRLVTAP TLRAQLAQDG VQLCALDDLD 300
SKRPGEVEPC EPTLEPFLEP SLL
  
```

Seq ID NO: 193 DNA sequence
Nucleic Acid Accession #: NM_005688.1
Coding sequence: 126..4439

80
85

```

1 11 21 31 41 51
CCGGGCAGGT GGCTCATGCT CGGAGCGGTG GTTGAGCGGC TGGCGCGGTT GTCTGGAGC 60
AGGGGCGCAG GAATCTGTAT TGAAACTTAA CAGTCTGTGA GCCCTGGAAC CTCGCTCAG 120
  
```


	AGAAGATGAA	GGATATCGAC	ATAGGAAAAG	AGTATATCAT	CCCCAGTCTT	GGGTATAGAA	180
	GTGTGAGGGA	GAGAAACGAG	ACTTCTGGGA	CGCACAGAGA	CGGTGAAGAT	TCCAAGTTCA	240
	GGAGAACTCG	ACCGTTGGAA	TGCCAAGATG	CCTTGGAAAC	AGCAGCCCGA	GCCGAGGGCC	300
5	TCCTCTTTGA	TGCCTTCAATG	CATTCTCAGC	TCAGAACTCT	GGATGAGGAG	CATCCCAAGG	360
	GAAAGTACCA	TCATGGCTTG	AGTGCTCTGA	AGCCCATCCG	GACTACTTCC	AAACACCAAGC	420
	ACCCAGTGGG	CAATGCTGGG	CTTTTCTCCT	GTATGACTTT	TTCTGTGGCTT	TCTTCTCTGG	480
	CCCGTGTGGG	CCACAAGAAG	GGGAGGCTCT	CAATGGAAGA	CGTGTGGTCT	CTGTCCCAAGC	540
	ACGAGTCTTC	TGACGTGAAC	TGCAGAAGAC	TAGAGAGACT	GTGGCAAGAA	GAGCTGAATG	600
10	AAGTTGGGCC	AGACGCTGCT	TCCCTGCGAA	GGTTGTGTG	GATCTTCTGC	CGCACCAAGC	660
	TCATCTGTGC	CATCGTGTGC	CTGATGATCA	CGCAGCTGGC	TGGCTTCAGT	GGACCAAGCT	720
	TCATGGTGAA	ACACCTCTTG	GAGTATACCC	AGGCAACAGA	GTCTAACCTG	CAGTACAGCT	780
	TGTTGTAGT	GCTGGGCTTC	CTCCTGACGG	AAATCGTGGG	GTCTTGGTGG	CTTGCACTGA	840
	CTTGGGCATT	GAATTACCGA	ACCGGTGTCC	GCTTGGCGGG	GGCCATCCTA	ACCATGGCAT	900
	TTAAGAAGAT	CCTTAAGTTA	AAGAACATTA	AAGAGAAATC	CCTGGGTGAG	CTCATCAACA	960
15	TTTGCTCAA	CGATGGGCG	AGAATGTTTG	AGGCAGCAGC	CGTTGGCAGC	CTGTGGGCTG	1020
	GAGGACCCGT	TGTTGCCATC	TTAGGCATGA	TTTATAATGT	AATTATTCTG	GGACCAACAG	1080
	GCTTCTGGG	ATCAGCTGTT	TTTATCCCTT	TTTACCCAGC	AATGATGTTT	GCATCAGGGC	1140
	TCACAGCATA	TTTCAGGAGA	AAATGCGTGG	CGGCCACGGA	TGAACGTGTC	CAGAAGATGA	1200
	ATGAAGTTCT	TACTTACATT	AAATTTATCA	AAATGTATGC	CTGGGTCAAA	GCATTTTCTC	1260
20	AGAGTGTTC	AAAAATCCCG	GAGGAGGAGC	GTCCGATATT	GGAAAAAGCC	GGGTACTTCC	1320
	AGGGTATCAC	TGTGGGTGGG	GCTCCCATTG	TGGTGGTGAT	TGCCAGCGTG	GTGACCTTCT	1380
	CGTTTCATAT	GACCCCTGGG	TTCCGATCTG	CAGCAGCACA	GGCTTTCACA	GTGGTGACAG	1440
	TCTTCAATT	CATGACTTTT	GCTTTGAAG	TAAACCCGTT	TTCACTAAAG	TCCCTCTCAG	1500
	AAGCTCAGT	AGCCTGTGAC	AGATTTAAGA	GTTTGTCTT	AATGGAAGAG	GTTCACATGA	1560
25	TAAAGAACA	ACCAGCCAGT	CCTCACATCA	AGATAGAGAT	GAAAAATGCC	ACCTTGGCAT	1620
	GGGACTCCTC	CCATCTCCAGT	ATCCAGAACT	CGCCCAAGCT	GACCCCAAAA	ATGAAAAAAG	1680
	ACAAGAGGGC	TCCAGGCGG	AAGAAAGAGA	AGGTGAGGCA	GCTGCAGCGC	ACTGAGCATC	1740
	AGGCGGTGCT	GGCAGAGCAG	AAAGGCCACC	TCTCCTGGA	CAGTGACGAG	CGGCCCAAGT	1800
30	CGAAGAGGA	AGAAAGGCA	CACATCCACC	TGGGCCACCT	GGCTTACAG	AGGACACTGC	1860
	ACAGCATCGA	TCTGGAGATC	CAAGAGGGTA	AATCTGCGGC	AGTGTGGGAA	1920	
	GTGGAAGAAC	CTCTCTCAAT	TCAGCCATTT	TAGGCCAGAT	GACGCTTCTA	GAGGCGCAGC	1980
	TTGCAATCAG	TGGAACCTCT	GCTTATGTGG	CCAGCAGGCG	CTGGATCCTC	AATGCTACTC	2040
	TGAGAGACAA	CATCCTGTTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
	ACAGCTGCTC	TCTGAGGCGT	GACCTGGCCA	TTCTTCCAG	CAGCGACCTG	ACGGAGATTG	2160
35	GAGAGCGAGG	AGCCAACTCG	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCTT	GCCCGGCGCT	2220
	TGTATAGTGA	CAGGAGCATG	TACATCTCTG	ACGACCCCTC	CAGTGCCTTA	GATGCCCATG	2280
	TGGCAACCA	CATCTTCAAT	AGTGCTATCC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
	TTGTTACCCA	CCAGTTACAG	TACCTGGTGG	ACTGTGATGA	AGTGATCTTC	ATGAAAGAGG	2400
	GCTGTATTAC	GGAAAGAGGC	ACCCATGAGG	AACTGATGAA	TTTAAATGGT	GACTATGCTA	2460
40	CCATTTTTAA	TAACTGTTTG	CTGGGAGAGA	CACCGCCAGT	TGAGATCAAT	TCAAAAAAGG	2520
	AAACCAAGTG	TTACAGAAAG	AAGTCACAAG	ACAAGGGTCC	TAAAAACAGG	TCAGTAAAGA	2580
	AGGAAAAGGC	AGTAAAGCCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
	GTTCAAGTGC	CTGGTCAGTA	TATGGTGTCT	ACATCCAGGC	TGCTGGGGGC	CCCTTGGCAT	2700
45	TCCTGTTTAT	TATGGCCCTT	TTTATGCTGA	ATGTAGGCG	CACCGCCTTC	AGCACCTGGT	2760
	GGTTGAGTTA	CTGGATCAAG	CAAGGAAGCG	GGAAACCCAC	TGTGACTCGA	GGGAACGAGA	2820
	CCTCGGTGAG	TGACAGCATG	AAGGACAATC	CTCATATGCA	GTACTATGCC	AGCATCTACG	2880
	CCCTCTCCAT	GGCAGTCATG	CTGATCCTGA	AAGCCATTCC	AGGAGTTGTC	TTTGTCAAGG	2940
	GCACGCTGGG	AGCTTCTCTC	CGGCTGCATG	ACGAGCTTTT	CCGAAGGATC	CTTGGAGGCC	3000
50	CTATGAAGTT	TTTGAACAG	ACCCCAACAG	GGAGGATTCT	CAACAGGTTT	TCCAAGAGCA	3060
	TGGATGAAGT	TGACGTGCGG	CTGCCGTTC	AGGCCGAGAT	GTTCATCCAG	AACTTATCC	3120
	TGGTGTCTTT	CTGTGTGGGA	ATGATCGCAG	GAGTCTTCCC	GTGGTTCCTT	GTGGCAGTGG	3180
	GGCCCCCTGT	CATCCTCTTT	TCAGTCTCTG	ACATTGTCTC	CAGGGTCTCT	ATTCGGGAGC	3240
	TGAAGCGTCT	GGACAATATC	ACGCAGTCAC	CTTCTCTCTC	CCACATCAGC	TCCAGCATAC	3300
55	AGGCCCTTGC	TACCATCCAC	GCCTACAATA	AAGGGCAGGA	GTTTCTGCAC	AGATACCAGG	3360
	AGCTGCTGGA	TGACAACCAA	GCTCCTTTTT	TTTGTGTTAC	GTGTGCGATG	CGGTGGCTGG	3420
	CTGTGCGGCT	GGACCTCATC	AGCATCGCCC	TCATCACCAC	CACGGGGCTG	ATGATCGTTC	3480
	TTATGACAGG	GCAGATTCCC	CCAGCCTATG	CGGGTCTGCG	CATCTCTTAT	GCTGTCCAGT	3540
	TAAAGGGGCT	GTTCAGATT	ACGGTCAGAC	TGGCATCTGA	GACAGAAGCT	CGATTCACTT	3600
60	CGGTGAGAG	GATCAATCAC	TACATTAAGA	CTCTGTCTCT	GGAAAGCACT	GCCAGAATTA	3660
	AGAACAAGGC	TCCCTCCCTC	GACTGGCCCC	AGGAGGGAGA	GGTGACCTTT	GAGAAGCGAG	3720
	AGATGAGGTA	CCGAGAAAC	CTCCCTCTTG	TCCTAAAGAA	AGTATCCTTC	ACGATCAAAC	3780
	CTAAAGAGAA	GATTGGCATT	GTGGGGCGGA	CAGGATCAGG	GAAGTCTCTG	CTGGGGATGG	3840
	CCCTCTTCCG	TCTGTGGGAG	TTATCTGGAG	GCTGCATCAA	GATTGATGGA	GAGGAATCA	3900
65	GTGATATTGG	CCTGTCCGAC	CTCCGAAGCA	AACCTCTAT	CATTCTCTAA	GAGCCGCTGC	3960
	TGTTCACTGG	CAGTGTGAGA	TCAAAATTGG	ACCCCTTCAA	CCAGTACACT	GAGAGCCAGA	4020
	TTTGGGATGC	CCTGGAGAGG	ACACACATGA	AAGAATGTAT	TGCTCAGCTA	CCTCTGAAAC	4080
	TTGAATCTGA	AGTGATGGAG	AATGGGGATA	ACTTCTCAGT	GGGGGAACGG	CAGCTCTTGT	4140
	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTAGATGAA	GCCACAGCTG	4200
70	CCATGGACAC	AGAGACGAC	TTATTGATT	AAGAGACCAT	CCGAGAAGCA	TTTGCAAGCT	4260
	GTACCATGCT	GACCATGTCC	CATCGCTTGC	ACACGGTCTT	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGCCCA	GGGACAGGTT	GTGGAGTTTG	ACACCCCATC	GGTCTTCTGT	TCCAACGACA	4380
	GTTCGCGATT	CTATGCCATG	TTTGTGCTG	CAGAGAACAA	GGTGGCTGTC	AAGGGCTGAC	4440
	TCCTCCCTGT	TGACGAAGTC	TCTTTCTTTT	AGAGCATTCG	CATTCCTGTC	CTGGGGCGGG	4500
75	CCCTCATAGT	CGTCTCTCTA	CCGAAACCTT	GCTTTCTG	ATTTTATCTT	TCGCACAGCA	4560
	GTTCGCGATT	GGCTTGTGTT	TTTCACTTTT	AGGGAGAGTC	ATATTTTGAT	TATTGTATT	4620
	ATTCCATATT	CATGTAAGCA	AAATTTAGTT	TTTGTCTT	ATTGCACTCT	AAAGAGTTCA	4680
	GGGAACCGTT	ATTATAATG	TATCAGAGGC	CTATAATGAA	GCTTTATACG	TGTAGCTATA	4740
	TCTATATATA	ATTCTGTACA	TAGCCTATAT	TTACAGTGAA	AATGTAAGCT	GTTTATTTTA	4800
80	TATTAATAA	AGACACTGTG	TAATAACAGT	GCATATTCTT	TTCTATCAIT	TTTGTACACT	4860
	TGCTGTACT	AGAGACTCTG	TTTGTCTATT	AGACTGTAGG	AAGAGTAGCA	TTTCAITCTT	4920
	CTCTAGCTGG	TGTTTTCAGG	GTGCCAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTTGGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCTCTGCCC	GCTTCCCTAC	AGCCGCTCCA	GGGGTGGGCT	5040
	GAGACGGGTT	GGCGGCTGGA	GACCATGACG	AGCGCCGTTA	GTTCTCAGGG	CTCTGCCTT	5100
85	CTGTCTGGTT	GTCATTTACT	GTTTCTGTCA	GGAGAGCAGC	GGGGCGAAGC	CCAGGCCCTT	5160
	TTTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTGCTCT	TCTTCTTTTT	GCTGTGTTT	CTAAACAGAA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCACTGCC	TCAGGTTCTT	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACCT	5340

GTTGGTCCCA AGCCCTGGAG CCAACTGCTG CTTTGTGAGG TGGCACTTTT TCATTGCTCT 5400
 ATTCCACAC TCACACAGT CAGTGGCAGG GCTCAGGATT TCGTGGGTCT GTTTCTCTTT 5460
 CTCACCGCAG TCGTGCACCA GTCTCTCTCT CTCTCTCCCC TCAAGTCTG CAACTTTAAG 5520
 CAGCTCTTGC TAATCAGTGT CTCACACTGG CGTAGAAGTT TTTGACTGT AAGAGACCT 5580
 ACCTCAGGTT GCTGTTTGTG GTGTGGTTTG GTGTGTTCCT GCAACCCCC TTTGTGCTGT 5640
 GGGGCTGGTA GCTCAGGTGG GCGTGGTCAC TGCTGTATC AGTTGAATGG TCAGCGTTGC 5700
 ATGTGCTGAC CAACTAGACA TTCTGTCCGC TTAGCATGTT TGCTGAACAC CTTGTGGAAG 5760
 CAAAATCTG AAAATGTGAA TAAATATT TTGATTTTG TAAAAAATA AAAAAAATA 5820
 AAAAAAATA AAAAAAATA

Seq ID NO: 194 Protein sequence:
 Protein Accession #: NP_005679.1

1 11 21 31 41 51
 MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRP TRPLECQDAL ETAARAEGLS 60
 LDASMSQLR ILDEEHPGKX YHGLSALKP IRTTSKHQHP VDNAGLFSCM TFSWLSLAR 120
 VAHKKGELSM EDVWSLSKHE SSDVNCRRLE RLWQELNEV GPDAASLRV VWIFCRTRI 180
 LSIVCLMITQ LAGFSGPAPM VKHLLEYTQA TESNLQYSL LVLGLLTLBI VRSWSLALTW 240
 ALNYRTGVR LRGAILTMAFK KILKLNKIE KSLGELINIC SNDGQRMFEA AAVGSLLAGG 300
 PVVAILGMIV NVILGPTGF LGSAPFIFY PAMMFASRLT AYFRRCVAA TDERVQKME 360
 VLTYYIKFIK YANVKAFQSQ VQKIREBER ILEKAGYFQG ITVGVAIVV VIASVVTFSV 420
 HMTLGFDLTA AQFTVTVTF NSMTFALKVT PFSVKSLSA SVAVDRFKSL FLMEEVHMIK 480
 NKPASPHIKI EMKNATLAWD SSHSIQNSP KLTPKMKDK RASRGKKEV RQLQRTHEQA 540
 VLAQKQHL LQSDERPSP EEEGKHILG HLRQLTLHS IDLEIQEGL VGICGSVSGG 600
 KTSLSIALI QMTLEGSIA ISGTFAYVAQ QANILNATLR DNILFKCYD EERYNSVLNS 660
 CCLRPDLAIL PSSDITEIGE RANLSCGQR QRISLARALY SDRSIYILD PLSALDAHV 720
 NHIFNSAIRK HLKSKTVLFV THQLQYLVD DEVIFMKEGC ITERGTHEEL MNLNGDYATI 780
 FNNLLGETP PVEINSKKEB SSGQKSKQD GPKTGSVKKE KAVKPEEQQL VQLEKQGQS 840
 VPWSYKGYI QAAGPLAFL VIMALFMLN GSTAFSTWVL SYNIKQSGN TTVTRGNETS 900
 VSDSMGNPH MQYASIALY SMAVMLILKA IRGVVFKGT LRASSRLHDE LFRILRLSPM 960
 KFFDTPTPGR ILARFSKMD EVDVRLPQA EMFIQNVILV FFCVGMAGV PFVFLVAVGP 1020
 LVILFSLVHI VSRVLIRELK RLDNITQSPF LSHITSSIQG LATIHAYNKG QEFLHRYQEL 1080
 LDDNQAPFFL FTCAMRWLAV RLDLISIALI TTTGLMIVLM HGQIPPAYAG LAISYAVQLT 1140
 GLFQPTVRLA SBTERTFSPV ERINHYIKTL SLEAPARIKN KAPSPDWQGE GEVTFENAEM 1200
 RYRENLPVLV KVSPTIKPK EKIGIVGRTG SGKSSLMAL FRLVELSGGC IKIDGVRISD 1260
 IGLADLRSL SIIPQEPVLF SGTVRSNLDP FNQYTEDQIW DALERTHME CIAQLPLKLE 1320
 SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETDLIIQE TIREAFADCT 1380
 MLTIAHRLHT VLGSDRIMVL AQGVVEFDT PSVLLSNDSS RPYAMFAAAE NKVAVRG

Seq ID NO: 195 DNA sequence
 Nucleic Acid Accession #: NM_006470
 Coding sequence: 228..1922

1 11 21 31 41 51
 GCTGCTCTGA GCCTGAGTAC TCTAGCTGCC TTGTGCGCAT CGCATCTGGC TGCCATCCAG 60
 CGCCAGCACA CAGTAATGAG TGGCCGAGCT TCCTCTGGGA GGGAGGAAAC AGTTAAATC 120
 TTGCAGCAGC TGCAATCATC TAGGCGTGTT TCTCTTGCT GACTTGGGCT GCACAGATCC 180
 TGGGCAAGG GACAGAGAA AGACAGCCTA GGAGCAGAGC CTCCAGATG GCTGAGTTGG 240
 ATCTAATGCG TCACGGGCCA CTGCCAGGG CCACTGCTCA GCGCCAGCC CCTCTCAGCC 300
 CAGACTCTGG GTACCCAGC CCAGATTCTG GGTACGCCAG CCACTGGAA GAAGAGGAGC 360
 TGGGCTCTCT GAGAGAGCTT GGCAGGGAGA CGGAGGAACA GGACAGCAGC TCTGCAGAGC 420
 AGGGGATTC TGTCTGTGAG GGGAAAGAGG TCTGTGTGA CTCTGCTCTT GATGACACCA 480
 GAAGAGTCAA GGCAGTGAAG TCTGTCTTAN CCGTGTGCTT GAATTAAGT GAAGAGCACT 540
 TGCAGCCGCA TCAGTGAAC ATCAAATGCT AAAGCCACCT GCTGACCGAG CCACTGAAGG 600
 ACCACAACCT GCGATACTGC CCGTCCACAC ACAGCCCACT GTCTGCTTTC TGCTGCCCTG 660
 ATCAGCAGTG CATCTGCCAG GACTGTGCTC AGGAGCAGAG TGGCCACACC ATAGTCTCCC 720
 TGGATGACAG CCGCAGGAGC AAGGAGGCTG AACTCCAGTG CACCCAGTTA GACTTGGAGC 780
 GGAACTCAA GTTGAATGAA AATGCCATCT CCAGGCTCCA GGCTAACCAA AAGTCTGTTG 840
 TGGTGTGCGT GTGAGAGTGC AAAGCGGTGG CTGAAATGCA GTTGGGGGAA CTCTGTGCTG 900
 CTGTGAGGAA GGGCCAGGCC AATGTGATGC TCTTCTTAGA GGAGAAGGAG CAAGCTGCGC 960
 TGAGCCAGGC CAAGGTATC AAGGCCACCC TGGAGTACAG GAGTGGCAG ATGAGAGAAG 1020
 CAAAGCAGGA GCTGAGAGG ATGGCGGCCA TCAGCAACAC TGTCCAGTTC TTGAGAGAGT 1080
 ACTGCAAGTT TAAGAACACT GAAGACATCA CCTTCCCTAG TGTTCAGTA GGGCTGAAG 1140
 ATAACTCTCT GGGCATCCGC AAGTTATCA CGGAATCCAC TGTACACTTA ATCCAGTTGC 1200
 TGGAGAACTA TAAGAAAAG CTCAGGAGT TTTCCAAGGA AGAGGAGTAT GACATCAGAA 1260
 CTCAAGTCTC TGCGTGTGTT CAGCGCAAT ATTGGACTTC CAAACCTGAG CCCAGCACCA 1320
 GGGAAACGTT CCTCCAATAT GCGTATGACA TCAGTTTGA CCGGACACA GCACACAAGT 1380
 ATCTCCGCTC CAGGAGGAGG AACCGCAAGG TCACCAACAC CAGCCCTGG GAGCATCCCT 1440
 ACCCGGACCT CCCAGCAGG TTCTGCACT GCGGCGAGGT GCTGTCCAG CAGAGTCTGT 1500
 ACCTGCACAG GTACTATTTT GAGGTGAGA TCTTCGGGGC AGGCACCTAT GTTGGCCTGA 1560
 OCTGCAAGG CATCGACCGG AAAGGGGAGG AGGCAACAG TTGCATTTC GGAACAACT 1620
 TCTCTGAGG CCTCCAATGG AACGGGAGG AGTTACCGG CTGGTACAGT GACATGGAGA 1680
 CCCCCTCAA AGCTGCGCTT TCCGGAGGC TCGGGTCTA TATCGACTTC CCGGAGGGA 1740
 TCTTTCTCTT CTATGCGTA GAGTATGATA CCATGACTCT GGTTCACAA TTTGCTGCA 1800
 AATTTTCAGA ACCAGTCTAT GCTGCTTCT GGCTTTCAA GAAGGAAAAC GCCATCCGGA 1860
 TTGTAGATCT GGGAGAGGAA CCGGAGAAC CAGCACGCTC CTTGGGGGTG ACTGCTCCCT 1920
 AGACTCCAGG AGCCATATCC CAGACCTTTG CCAGTACAG TGTAGGGATT TGCATTTTAG 1980
 GGTGATTGTT GGCAGAAAT AACTGCTGAT GGTAGCTGGC TTTTGAATC CTATGGGGTC 2040
 TCTGAATGAA AACATTCTCC AGCTGCTCTC TTTTGTCCA TATGGTGTG TTCTCTATGT 2100
 GTTTCAGTA ATTCTTTTTT TTTTTTTGA GACGGAGTCT GCACTGTGG CCCAGGCTGG 2160
 AGAGCAGTGG CCGATCTCTG GCTCACTGCA AGCTCCGCT CCGAGTTCA AGCAATCTC 2220
 CTGCTCAGC CTCCGAGTA GCTGGGATTA CAGTGCCTG CCACCAACC CAGCTAATGT 2280
 TTTGATTTT TAGTAGAGAT GGGGTTTCAC CATGTTGGCC AGGCAGATCT CAACTCCTG 2340

ACCTGCTGAT GCACCCACCT CGGCTCCCA AAGTGTCTGG ATTACATGCG TGAGCCACTG 2400
 CGCCCTGCTT GTTTGTAGTA ATTTTATAGG ACCAAATCTC CCTCATCTTC TAGTGCCTAT 2460
 CTCCTCTCTG TTAAGGTAAAT TGTCACTCTG TGCCCGAAGT GGATGACGAC GAACCTTAAAT 2520
 GAGTGGCTGA AAAGATTGCA GAGTTATCAT AATAAATGTC TAACTTGGCT

Seq ID NO: 196 Protein sequence:
 Protein Accession #: NP_006461

1 11 21 31 41 51
 MAELDLMAPG PLPRATAQPP APLSPDSGSP SPDSGSASPV EEDVGSSEK LGRETEEQDS 60
 DSAEQGDPAQ EGKEVLCDFC LDDTRRVKAV KSLTCTMVNY CEEHLQPHQV NIKLQSHLLT 120
 EPVIDENWRY CPAHHSPLSA FCCPDQCCIC QDCQEHSGH TIVSLDAARR DKEAELQCTQ 180
 LDLERKLKLN ENAISRLQAN QKSVLVSVSE VKAVAEMQFG ELLAAVRKAQ ANVMLFLEEK 240
 EQAALQANG IKAHLEYRSA EMEKSKQELE RMAAISNTVQ FLEEYCKFKN TEDITFPSVY 300
 VGLDKLSGI RKVITESTVH LIQLLENYKK KLQEFSEKEE YDIRTQVSAV VQRKYWTSKP 360
 EPSTRBQFLQ YAYDITFDDP TAHKYLRLQE ENRKVTNTTP WEHPYPDLPS RFLHWRQVLS 420
 QQSLYLHRYF FEVRIFGAGT VVGLTCKGID RKGEERNSCI SGNFWSLSQ WNGKEFTAWY 480
 SDMETPLKAG PFRRLGVYID FPGGILSPYG VEYDTMTLVH KFACKFSEPV YAAFWLSKKE 540
 NAIRIVDLGE EPEKPAPSLG VTAP

Seq ID NO: 197 DNA sequence
 Nucleic Acid Accession #: NM_004316
 Coding sequence: 433-1149

1 11 21 31 41 51
 CCCGAGACCC GGGCGAAGAG AGCGCAGCCT TAGTAGGAGA GGAACGCGAG ACGCGGCAGA 60
 GCGCGTTCAG CACTGACTTT TGCTGTGCT TCTGCTTTTT TTTTCTTAG AAACAAGAAG 120
 GCGCCAGCGG CAGCCTCACA CGCGAGCGCC ACGCGAGGCT CCGGAAGCCA ACCCGCGAAG 180
 GGAGAGGGGG AGGGAGGAGG AGGCGGCGTG CAGGGAGGAG AAAAAGCATT TTCACCTTTT 240
 TTGCTCCAC TCTAAGAAGT CTCGCGGGA TTTTGTATAT ATTTTAAAC TTCGCTCAGG 300
 GCTCCCGCTT CATATTTCTT TTTCTTTCCC TCTCTGTTC TGCACCCAAG TTCCTCTCTG 360
 GTCCCTCTCG CGGGCCCCGC ACCTCGCGTC CCGGATCGCT CTGATTCCGC GACTCCTTGG 420
 CCGCGCTGCG GCATGGAAGG CTCTGCCAAG ATGGAGAGCG GCGCGCGCGG CCAGCAGCCC 480
 CAGCCGACAG CCGCAGCAGC CTTCCTGCGG CCGCAGCCT GTTCTTTGC CACGGCCGCA 540
 GCGCGGCGGG CCGCAGCGCG CGCAGCGCGA GCGCAGAGCG CGCAGCAGCA GCAGCAGCAG 600
 CAGCAGCAGC AGCAGCAGCA GCAGGCGCGG CAGCTGAGAC CGGCGGCGGA CGGCGAGCCC 660
 TCAGGGGGCG GTACAAAGTC AGCGCCCAAG CAAGTCAAGC GACAGCGCTC GTCTTCGCCC 720
 GAAGTATGAC GCTGCAAAAG CCGGCTCAAC TTCAGCGGCT TTGGCTACAG CCTGCCCGAG 780
 CAGCAGCGCG CGCGCTGTGC GCGCGCGAAG GAGCGCGAGC GCAACCGCGT CAAGTTGGTC 840
 AACCTGGGCT TTGCCACCTT TCGGGAGCAC GTCCCCAAGC GCGCGGCCAA CAAGAAGATG 900
 AGTAAGGTGG AGACACTGCG CTCGCGGTC GAGTACATCC GCGCGCTGCA GCAGCTGCTG 960
 GACGAGCATG ACCCGGTGAG CGCGGCTTC CAGGAGGCGG TCCTGTGCGC CACCATCTCC 1020
 CCCAAGTACT CCAAGCACTT GAAGTCCATG GCGCGCTGCG CGGTCTCATC CTACTGTCG 1080
 GAGGAGGGCT CTACGACCCC GCTCAGCCCC GAGGAGCAGG AGCTTCTCGA CTTCAACAAC 1140
 TGGTCTGAG GCGCTCGGCC TGGTCAGGCC CTGGTGCGAA TGGACTTTGG AAGCAGGGTG 1200
 ATCGCAACA CTGATCTTTT AGTCCTTTCT TGTCACTGCG GTTGGGAGGG GGAGAAAAGG 1260
 AAAAGAAAAA AAAAGAAAGA GAAGAAGAAA AGAGAAGAAG AAAAAAACGA AAACAGTCAA 1320
 CCAACCCCAT CGCAACTATA GCGAGGCATG CCTGAGAGAC ATGGCTTTCA GAAACCGGGA 1380
 AGCGCTCAGA ACAATATCTT TGCACTCCAA TCATTACCG AGATATGAAG AGCAACTGGG 1440
 ACCTAGTATA ATCGCATAAA TGCAGCTTGT GTGCAAAAGC AGTGGGCTCC TGGCAGAAAG 1500
 GAGCAGCACA CGCGTTATAG TAAGTCCCAT CACCTCTAAC ACGCACAGCT GAAAGTTCTT 1560
 GCTCGGGTCC CTTCACCTCC CGGCCCTTTC TTAGAGTGCA GTTCTTAGCC CTCTAGAAAC 1620
 GAGTTGGTGT CTTTC

Seq ID NO: 198 Protein sequence:
 Protein Accession #: NP_004307

1 11 21 31 41 51
 MESSAKMESG GAGQQPQPQP QQPFLPPAAC FFATAAAAAA AAAAAAQA QQQQQQQQQQ 60
 QQQQAPQLRP AADGQPSGGG HKSAPKQVKR QRSSSPELMR CKRRLNFSGF GYSLPQQQPA 120
 AVARRNERER NRKVLNLGF ATLREHVPNG AANKMKSKE TLRSAVEYIR ALQQLLDEHD 180
 AVSAAFQAGV LSPTISPNYS NDLNSMAGSP VSSYSSDEGS YDPLSPEEQE LLDFTNWF

Seq ID NO: 199 DNA sequence
 Nucleic Acid Accession #: NM_007015
 Coding sequence: 1-1005

1 11 21 31 41 51
 ATGACAGAGA ACTCCGACAA AGTTCCTCAT GCCCTGGTGG GACCTGATGA CGTGGAAATC 60
 TGACAGCCCC CGCGCTAGCG TACGCTGACG GTGAAGCCCT CCAGCCCCCG GCGGCTGTCT 120
 AAGGTGGAG CGGTGGTCTT CATTTGGGGA GCTGTGCTGC TGCTCTTTGG GGCCATCGGG 180
 GCCTTCTACT TCTGGAAGGG GAGCGACAGT CACATTACA ATGTCCATTA CACCATGAGT 240
 ATCAATGGGA AACTACAAGA TGGGTCAATG GAAATAGAG CTGGGAACAA CTGGGAGACC 300
 TTTAAATGG GAAGTGGAGC TGAAGAAGCA ATTGCAGTTA ATGATTTCGA GAATGGCATC 360
 ACAGGAATTC GTTTGTCTGG AGGAGAGAGG TGCTACATTA AAGCGCAAGT GAAGGCTCGT 420
 ATTCTGAGG TGGGCGCGGT GACCAACAG AGCATCTCCT CCAACTGGA AGGCAAGATC 480
 ATGCCAGTCA AATATGAAGA AAATTCTCTT ATCTGGGTGG CTGTAGATCA GCCTGTGAAG 540
 GACAACAGCT TCTTGAGTTC TAAGGTGTGA GAAGTCTGCG GTGACCTTCC TATTTCTGCG 600
 CTTAAACCAA CTTATCCAAA AGAAATCCAG AGGGAAGAG GAGAAGTGGT AAGAAAAATT 660
 GTTCCAACTA CCACAAAAG ACCACACAGT GGACCAAGCA GCAACCCAGG CGCTGGAAGA 720
 CTGAATAATG AAACAGACCC CAGTGTTCAG GAGGACTCAC AAGCCTTCAA TCCTGATAAT 780

5
10
15
20
25
30
35
40
45
50
55
60
65

CCTTATCATC AGCAGGAAGG GGAAAGCATG ACATTGACC CTAGACTGGA TCAAGGAAGG 840
ATCTGTTGTA TAGAATGTAG GCGAGAGCTAC ACCCACTGCC AGAAGATCTG TGAACCCCTG 900
GGGGGCTATT ACCCATGGCC TTATAATTAT CAAGGCTGCC GTTCGGCCTG CAGAGTCATC 960
ATGCCATGTA GCTGGTGGGT GGCCCGTATC TTGGGCATGG TGTGAATCA CTTTATATAT 1020
CACGTGCTGT AAAATAAGAA CTAGCTGAAG AGACAACCAA AGAAGCATT AAGCAGGTG 1080
ATGCTGATGG GACCATAAAA TATTTTACA CGCAGCCTGA GCGGTATTTC TTGACACTCT 1140
TAACAGAATT TTTTAAATCG TTTTCCAGAA CTTTAGTATA TGCAAATGCA CTGAAAGGGT 1200
AGTTCAAGTC TAAATGCCA TAACCCCGTT ATTGTTATT TTTTATTGTC ATTGATTGTC 1260
CATAAGTCTT CCCTTGCTTG CATCTTCCAA AGCTATTTTG AAATAAACAC GAAATTTTAC 1320
AGTTTGCC

Seq ID NO: 200 Protein sequence:
Protein Accession #: NP_008946

1 11 21 31 41 51
MTENS DKVPI ALVGPDDVEF CSPPAYATLT VKPSSPARLL KVGAVVLISG AVLLLFPAIG 60
AFYFWKSDS HIYNVHYTMS INGKLQDGSN EIDAGNNLET FMGSGAEEA IAVNDFQNGI 120
TGIRFAGGEK CYIRAKQVKR IPEVGAVTKQ SISKLEBKI MPVKEYEENL IWWAVDQPVK 180
DNSFLSSKVL ELCGDLPIFW LKPTYPKEIQ RERREVRKI VPTTKRPHS GPRSNPGAGR 240
LNNETRFVSQ EDSQAFNPDN PYHQEGESM TFDPRLDHEG ICCIECRSSV THCQKICEPL 300
GGYYPWFYNY QGCRSACRVI MPCSWWVARI LGMV

Seq ID NO: 201 DNA sequence
Nucleic Acid Accession #: NM_000728.2
Coding sequence: 112..495

1 11 21 31 41 51
GTAATAAGAG CGGGGTCTCC GCGGGGAAGG CGCCACAGC AGGTGTGGTG TTCATCCCGG 60
GTCGACCGGC CGCTCGCGCT GCGCTGAAAC TCTAGTGGCC AGAGAGCGCG CATGGGTTTC 120
CGGAAGTTCT CCCCTTCTCT GGCTCTCAGT ATCTGTTGCC TGTACCGGCC GGGCAGCCTC 180
CAGGCGCGGC CATTCAGGTC TGCCCTGGAG AGCAGCCAG ACCCGGCCAC ACTCAGTAAA 240
GAGGACGCGC GCCTCTGCTG GGCTGCACTG GTGCAGGACT ATGTGCAGAT GAAGGCCAGT 300
GAGCTGAGGC AGGAGCAGGA GACACAGGCG TCACGCTCCG CTGCCCGAAG GAGAGCCTGC 360
AACACTGCCA CCTGTGTGAC TCATCGGCTG GCAGGCTTGC TGAGCAGATC AGGGGGCATG 420
GTGAAGAGCA ACTTCGTGCC CACCAATGTG GGTTCCAAAG CCTTTGGCAG GCGCGCAGG 480
GACCTTCAAG CCTGAGCAGA TGAATGACTC CAGGAAGAAG GTGTGTCTTA AATCCAATGA 540
CATATCCCTA TAAGAGATTC ACTCAGAAGA CACATGTGGA GAAGGTGACA TGACAGAGGC 600
AAGGAGGCAC AAGCCAAGCA AGCTGTGTGC TACCAGAAGC CAGAAATACA GAACAGTCTC 660
TGGAAGAAGA GCAGCCCTGC TGACACCTAG AGTTTGGACT TCACGCTTCC AGAAGCTGTG 720
GAGAATAATT TCTGTTGTTT TAAGCCACAA AGTTTGTGTT AATTGTTTAT GACAGCCCTA 780
GGAAACTAAT ACAATACATT TTCAATTTAT TTGGGTAAT GCCTTGGAGT GGGATTGCTG 840
GGTTATTGG AAGTGTGTA TTTAACTCTG TAAGAACTCT CCAAACTATT TTCTGAAGTG 900
ACTGTACCAC TTCGCTTCTT TGCCAGCCAC ATATGAGAGC TCTAGTATT TCCCAAAATAG 960
GTATGTAGCA GTATCTCATT GCTGTTTAA TTTGTATTTC CCAATGACT AATGACGTTG 1020
AGCATCTATT TTACCATATG TTTATCACCT TTATGGAAG GTCTGTTTAA ATCTCTGCT 1080
AAATTTTGT TGGCTTGCTT GCTTTATTAG TGTGTAGTTT TTAGAGCTCT TTATATGTTG 1140
TGGATGCAAG ATTGTTTTC GATATATAGT TTGGAACCTT CCTTCCCTG AATCTGCGGA 1200
TTGCTTTTTC ATTTCTTAGT CAGTGTCTCT CACAGAGAAA AAGTTGTAAT TTGAATAAGA 1260
TCCAATTCAT CTTTTTTTTT CTTTATGTA TTGTGCTTTT AGTTCAITGTC TAAGAACTCT 1320
TTGCCAATC AAGGTCCCAA GGTCAACAATA ACCTTATCT ATACTTCTT GTAAAAAGTT 1380
TATAGTTTGA TATTTTATAT GTAGATTAGT GATCTATTTT GAGTTAATTT TTGTATAAG 1440
TGAGAGGTGT AGGTTGAAAT TCATACCTGT GAATATAGAT ACCCAATTGT TTCAGTGCCA 1500
TTTGTAAAAA AGACTGTAT TTCAACATTT AATGCCCCCT GCACCTTGT CAAAAAGCAA 1560
CTGATCATAT TTGTGTGGGT ATATTTCTGG GTTCTCAATT CTGCTCATT GATTGATTG 1620
ACCATCTCTT TGCCCAATGT ATACTGCCTT GATTAGTGA GTGTAAAGT GAATCTCAA 1680
ACCAATAAT GTGGGTCTAC CAACATTGTT CATTCTGTT CAAAAAGAT TTAGCTACAT 1740
CTAAAAATAT TTCTACATCT TTTATACATT TTGAATCAG TGTGTACTA TCTACAAAAT 1800
TTCTGATGAG ATTTTAAATG GGATTGTGTT AAATCAGTGG GTTAATTTTG GGGAATTAG 1860
CATATTAATA ATATTAAGTC GTTCAATTCA TGAACACAAT ACATGTTTTC ACTTATTAG 1920
GTTTTCTCTG TTTTCTTTT TTTAACAGTG TTCTCAGTT TCAACAGAAA TATTCTACAC 1980
ATATCTGTT AGATTTTAA CTATTTTATT TTTTGGTGCT AATGTAAATG GTACTTAAAC 2040
ATTTTGTGTT TTAATTGTTT ATTGCTAGTA GATAGAAATA CAATATTTAA AATATTAGGA 2100
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 202 Protein sequence:
Protein Accession #: NP_000719.1

1 11 21 31 41 51
MGPRKPSFPL ALSILVLYQA GSLQAAPFRS ALESSPDPAT LSKEDARLLL AALVQDYVQM 60
KASELKQEQE TQSSSSAAQK RACNTATCVT HRLAGLLSRG GGMVKSFNVP TNVGSKAPGR 120
RRRDLQA

Seq ID NO: 203 DNA sequence
Nucleic Acid Accession #: NM_001741
Coding sequence: 71..496

1 11 21 31 41 51
CTCTGGCTGG ACGCGCGCGC CGCCGCTGCC ACCGCTCTG ATCCAAGCCA CCTCCCGCCA 60
GAGAGGTGTC ATGGGCTTCC AAAAGTTCTC CCCCTTCTG GCTCTCAGCA TCTTGGTCTC 120
GTTGCAGGCA GCGACGCTCC ATGCAGCACC ATTCAGGTCT GCGCTGAGA GCAGCCAGC 180
AGACCCGCGC AGCCTCAGTG AGGACGAAGC GCGCCTCTG CTGGCTGCAC TGGTGACAGG 240
CTATGTGCAG ATGAAGGCCA GTGAGCTGGA GCAGGAGCAA GAGAGAGAGG GCTCCAGCCT 300

5 GGACAGCCCC AGATCTAAGC GGTGCGGTAA TCTGAGTACT TGCATGCTGG GCACATACAC 360
 GCAGGACTTC AACAGTTTC ACACGTTCOC CCAAAGTGCA ATTGGGGTTG GAGCACCTGG 420
 AAAGAAAAGG GATATGTCCA GCGACTTGGA GAGAGACCAT CGCCCTCATG TTAGCATGCC 480
 CCAGAAATGCC AACTAAACTC CTCCTTTTCC TTCTTAATTT CCCTTCTTGC ATCCTTCTTA 540
 TAACCTGATG CATGTGGTGT GGTTCCTCTC TGGTGGCTGT TTGGGCTGGT ATTGGTGGCT 600
 TTCCTTGTGG CAGAGGATGT CTCAAACTTC AGATGGGAGG AAAGAGAGCA GGACTCACAG 660
 GTTGAAGAG AATCACCTGG GAAAATACCA GAAAATGAGG GCGCTTTGA GTCCCCAGA 720
 GATGTCATCA GAGCTCTCTC GTCTGCTTC TGAATGTGCT GATCATTGGA GGAATAAAAT 780
 TATTTTCC C

Seq ID NO: 204 Protein sequence:
 Protein Accession #: NP_001732

15 1 11 21 31 41 51
 MGFKQFSPFL ALSILVLLQA GSLHAAPFRS ALESSPADPA TLESEARLL LAALVQDYVQ 60
 MKASELEQEQ EREGSSLDSP RSKRCGNLST CMLGTYTQDF NKFTFPQTA IGVGAPGKKR 120
 DMSDLERDH' RPHVSMQNA' N

Seq ID NO: 205 DNA sequence
 Nucleic Acid Accession #: NM_005361
 Coding sequence: 1-945

25 1 11 21 31 41 51
 ATGCCTCTTG AGCAGAGGAG TCAGCACTGC AAGCCTGAAG AAGGCCTTGA GGCCCGAGGA 60
 GAGGCCCTGG GCGCTGGTGGG TGCGCAGGCT CCTGCTACTG AGGAGCAGCA GACCGCTTCT 120
 TCCTCTCTTA CTCTAGTGGG AGTTACCTCG GGGGAGGTGC CTGCTGCCGA CTACCCGAGT 180
 30 CACTCCCAACA CTCCTCAGGG AGCCTCCAGC TTCTCGACTA CCATCAACTA CACTCTTTGG 240
 AGACAATCCG ATGAGGGGTC CAGCAACCAA GAAGAGGAGG GGCCAAGAAT GTTCCCGGAC 300
 CTGAGTCCG AGTTCCCAAG AGCAATCAGT AGGAAGATGG TTGAGTTGGT TCATTTTCTG 360
 CTCCTCAAGT ATCAGAGCCAG GGAGCCGGTC ACAAAGCCAG AAATGCTGGA GAGTGTCTCT 420
 AGAAATGGCC AGGACTTCTT TCCCGTGATC TTCAGCAAGG CCTCGAGTA CTTCGAGCTG 480
 35 GTCTTTGGCA TCGAGGTGGT GGAAGTGGTC CCCATCAGCC ACTGTACAT CTGTGTCAAC 540
 TGCCGTGGCC TCTCTACGA TGGCTGCTG GGCACATC AGGTGATGCC CAAGACAGGC 600
 CTCCTGATAA TGTCTCTGGC CATAATCGCA ATAGAGGGCG ACTGTGCCCC TGAGGAGAAA 660
 ATCTGGGAGG AGCTGAGTAT GTTGAGGTG TTTGAGGGGA GGGAGGACAG TGTCTTCGCA 720
 CATCCAGGGA AGCTGCTCAT GCAAGATCTG GTGCAGGAAA ACTACCTGGA GTACCGGCAG 780
 40 GTGCCCGGCA GTGATCTGTC ATGCTACGAG TTCTGTGGG GTCCAAGGGC CCTCATTGAA 840
 ACCAGCTATG TGAAGTCTCT GCACCATACA CTAAAGATCG GTGGAGAACC TCACATTTC 900
 TACCCACCCC TGCATGAACG GGCTTTGAGA GAGGAGAGAG AGTGA

Seq ID NO: 206 Protein sequence:
 Protein Accession #: NP_005352

45 1 11 21 31 41 51
 MPLEQRSQHC KPEBGLAARG EALGLVGAQA PATEEQQTAS SSSTLVEVTL GEVPAADSPS 60
 50 PPHSPQGASS PSTTINYTLW RQSDGSSNQ EEEGRPMFPD LESEFQAAS RKMVELVHFL 120
 LLKYRAREPV TKAEMLESVL RNCQDFPFI FSKASEYLQL VFGIEVVEVV PISHLYILVT 180
 CLGLSYDGLL GDNQVMPKTC LLIIVLAIIA IEGDCAPEEK IWEELSMLEV FEGREDSVFA 240
 HPRKLLMQDL VQENYLEYRQ VPGSDPACYE FLWGPRALIE TSYVKVLHHT LKIGGEPHIS 300
 YPPLHERALR EGEE

Seq ID NO: 207 DNA sequence
 Nucleic Acid Accession #: NM_021115
 Coding sequence: 743-2893

60 1 11 21 31 41 51
 AAAGGAAGGG AGGGAGGGAG AAAGGAGAAG TTGGTTTAGA GGCCAGCCGG ACGAGCTTTG 60
 GGCACCGCCC TTAGGAGGGC CACCTCAGA GTCTGACAGC AGGTGAAGGT CCTAAATCTC 120
 65 CCCAACTAA CTGGTGTCTT TTCTCTCTT CCAAGATGCT CTTCCTCGAG GAGATGCTAG 180
 CCCTTTGGGT CCTTACCTCC TGCCCTCAGG AGCCCGGAG AGAGGCAGTC CTGGCAAAGA 240
 GCACCTGTA GAGAGAGTGG TAACAGCGCC CCCAGTTCC TCACAGTCGG CGGAAGTGCT 300
 GGGCGAGCTG GTGCTGGATG GGACCGCACC CTCTGCATC CACGACATCC CAGCCCTGTC 360
 ACCGCTGCTT CCAAGAGGAG CCCGCCCAA GCACGCTTG CCCCCAAGA AGAAACTGCC 420
 TTGCTCAAG CAGGTGAAC CTGCCAGGAA GCAGCTGAGG CCCAAGGCC CCTCCGAGC 480
 70 CACTGTCAA AGGGCAGGGT CCCAGCCAGC GTCCAGGGC CTAGATCTCC TCTCTCTC 540
 CACGAGAGG CTTGGCCAC CGGGGAGCCC GGACCCATC GTGGCTCCG AGGAGGCATC 600
 AGAAGTGCCT CTTTGGCTGG ACCGAAAGGA GAGTGGGTC CACTAACAC CCGCACCCCT 660
 GCAATCTCC CCTTCACTT CCGAGCCCTA TGTGGCCAC AACTTCCCC AGAGGCCAGA 720
 ACCCGGGGAG CTTGGGCTG ACATGGCCCA GAGGCCCCC CAGGAGGACA CCAGCCCCAT 780
 75 GGCCTGATG GACAAAGGTG AGAATGAGCT GACTGGGTCA GCCTCAGAGG AGAGCCAGGA 840
 GACCACTACC TCCACATTA TCACACCCAC GGTGATCACC ACCGAGCAGG CACCACTCT 900
 CTGAGTGTG AGCTTCTCCA ATCTGAGGG GTACATTGAC TCCAGCACT ACCCACTGCT 960
 GCCCTCAAC AACTTCTCG AGTGACATA CAACGTGACA GTCTACACTG GCTATGGGGT 1020
 GGAGCTCCAG GTGAAGAGTG TGAACCTGTC CGATGGGGAA CTGCTCTCCA TCCCGGGGT 1080
 80 GGAAGGCGCT ACCCTGACCG TCTGGCCAA CCAACACTC CTGGTGGAG GGCAGGTAAT 1140
 CCGAAGCCCC ACCAACCACT TCTCGTCTA CTTCGAGACC TTCCAGGAGC ACGGCTTGG 1200
 GACCTTCCAG CTTCACTAT AGGCCCTCAT GCTGAGCTGC AACTTCCCC GCGGCGCTGA 1260
 CTCGGGGAT GTACCGGTGA TGGACCTGCA CTCAGGTGGG GTGGCCACT TTCCTGCCA 1320
 85 CCTGGGCTAT GAGCTCCAGG GCGCTAAGAT GCTGACATGC ATCAATGCTT CCAAGCCGCA 1380
 CTGGAGCAGC TCTGCTCAGC TCTTGTGGA GGGGCAGTGC ACAATGCCAC 1440
 CATGGGCGC GTCTCTCCC CAAGTTACCC TGAACACACA AATGGAGGCC AATTCTGCAT 1500
 CTGGAGGATT GAAGCTCCAG AGGGCCAGAA GCTGCACCTG CACTTTGAGA GCGCTGTGCT 1560

5
10
15
20
25

```

GCATGACAAG GACAGGATGA CGGTTACAG CGGGCAGACC AACAGTCAG CTCCTCTCTA 1620
CGACTCCCTT CAAACCGAGA GTGTCCCTTT TGAGGGCCTG CTGAGCGAAG GCAACACCAT 1680
CGGCATCGAG TTCAGTCCCG ACCAGGCCCG GGGCGCCTCC ACCTTCAACA TCCGATTTGA 1740
AGCGTTTGGAG AAAGGCCACT GCTATGAGCC CTACATCCAG AATGGGAAT TCACTACATC 1800
CGACCCGACC TATAACATTG GGAATATAGT GGAGTTCACC TGGGACCCCG GCCACTCCCT 1860
GGAGCAGGGC CCGGCCATCA TCGAATGCAT CAATGTGGGG GACCCATACT GGAATGACAC 1920
AGAGCCCGCT TGCAGAGCCA TGTGTGGTGG GGAGCTCTCT GCTGTGGCTG GGGTGGTATT 1980
GTCCCAAAAC TGGCCCGAGC CTAAGTGGGA AGGTGAAGAT TGTATCTGGA AGATCCAGCT 2040
GGGAGAAGAG AAACGGATCT TCTTAGATAT CCACTTCCTG AATCTGAGCA ACAGTGACAT 2100
CTTGACCATC TACGATGGCG AGGAGGTCAT GCCCCACATC TTGGGGCAGT ACCTTGGGAA 2160
CAGTGGCCCC CAGAAATCTG ACTCCTCCAC GCCAGACTTA ACCATCCAGT TCCATTGGGA 2220
CCCTGCTGGC CTCATCTTTG GAAAGGGCCA GGGATTATC ATGAACATA TAGAGGTATC 2280
AAGGAATGAC TCCTGCTCGG ATTTACCCGA GATCCAGAAT GGCTGGAAAA CCACTTCTCA 2340
CACGGAGTTG GTGCGGGGAG CCAGAATCAC CTACCAGTGT GACCCCGGCT ATGACATCGT 2400
GGGGAGTGAC ACCCTCACCT GCCAGTGGGA CCTCAGCTGG AGCAGCGACC CCCCATTITG 2460
TGAGAAAATT ATGACTGCA CCGACCCCGG AGAGGTGGAT CACTCGAGCC GCTTAATTTT 2520
GGATCTCTGT CTGCTGGTGG GGACCAACAT CCAATACACC TGCAACCCCG GTTTTGTGCT 2580
TGAAGGAGAT TCTCTCTGA CCTGCTACAG CCGTGAACA GGGACTCCCA TCTGGACGTC 2640
TCGCCTGCC CACTGCGTTG CAGAAGCGGC AGCAGAGACG TCGCTGGAAG GGGGGAACAT 2700
GGCCCTGGCT ATCTTCATCC CGGTCTCAT CATCTCCTTA CTGCTGGGAG GAGCCTACAT 2760
TTACATCACA AGATGTCTGT ACTATTCCAA CCTCCGCTCT CCTCTGATGT ACTCCACCC 2820
CTACAGCCAG ATCAGCGTGG AAACCGAGTT TGACAACCCC ATTTACGAGA CAGGGGGAAC 2880
CCAAAAGGTT TAGGGTTTCA TTTAAAAGA GGTACCCCTT AAAAAGGGGC TTGTGAACCT 2940
AACCCTAATT TCCCCGAGAC ATTTATCCAA AGGCCCTGGG GGCCTTGATT TAAACCCCA 3000
AAAGGCGGCT GTTTTTTGGT TAAACTTTTT AACAAAGGT TACGGGTTTT TTCCCGGAT 3060
TTTATAAATT TTAAGAATG

```

30 Seq ID NO: 208 Protein sequence:
Protein Accession #: NP_065938

35
40
45

```

1 11 21 31 41 51
MAQEAPEQEDT SPMALMDKGE NELTGSASEE SQETTTSTII TTTVITTEQA PALCSVSFSN 60
PEGYIDSSDY PLLPLANFLE CTYNVTVTYTG YGVELQVKS NLSGZELLSI RGVDPGLTV 120
LANQTLLEVG QVIRSPNTI SVYFRTFQDD GLGTFQLHYQ AFMLSCNFPR RPSGDVTVM 180
DLHSGGVAHF HCHLGYELQG AKMLTCINAS KPHWSSQEPF CSAPCGGAVH NATIGRVLSP 240
SYPPENTNGSQ FCIWTIEAPE GQKLHLHFER LLLHDKDRMT VHSQGTNKSA LLYDSLQTES 300
VPFEGLLSEB NTIRIEFTSD QARAASTFNI RFEAFKGHG YEPYIQNGNF TSDPTYNIG 360
TIVEPTCDPG HSLEQGPAIL ECINVRDPYH NDTEPLCRAM CGGELSAVAG VVLSFNPPEP 420
YVEGEDCIWK IHVGEERIF LDIQFLNLSN SDILTIYDGD EVMPHILGQY LGNSGPQKLY 480
SSTPDLTIQF HSDPAGLIFG KGQGFIMNYI EVSRNDSGSD LPEIQNGWKT TSHTLVIRGA 540
RTTYQCDPGY DIVGSDTLTC QWDLWSWSDP PFCEKIMYCT DPEVDHSTR LISDPVLLVG 600
TTIYQTCNPG FVLGSSLLT CYSRETGTPI WTSRLPHCVS EAAETSLEG GNMALAIPIP 660
VLIISLLGG AYIYITRCRY YSNLRLPLMY SHPYSQITVE TEFDNPIYET GGTQKV

```

50 Seq ID NO: 209 DNA sequence
Nucleic Acid Accession #: NM_001327.1
Coding sequence: 89-631

55
60
65

```

1 11 21 31 41 51
AGCAGGGGGC GCTGTGTGTA CCGAGAATAC GAGAATACCT CGTGGGCCCT GACCTTCTCT 60
CTGAGAGCCG GGCAGAGGCT CCGAGGCCAT GCAGGCCGAA GGCAGGGGCA CAGGGGGTTC 120
GACGGGGCGAT GCTGATGGCC CAGGAGGCCC TGGCATTCTT GATGGCCAG GGGGCAATGC 180
TGGCGGCCCA GGAGAGCGGG GTGCCACGGG CGGCAGAGGT CCGCGGGGCG CAGGGGCGAG 240
AAGGGCTCTG GGGCCGGGAG GAGGCGCCCC GCGGGGTCCG CATGGCGGCG CGGCTTCAGG 300
GCTGAATGGA TGCTGCAGAT GCGGGGCCAG GGGGCCGGAG AGCCGCTCTG TTGAGTTCTA 360
CCTCGCCATG CCTTTCCGCA CACCCATGGA AGCAGAGCTG GCCCGCAGGA GCCTGGCCCA 420
GGATGCCCCA CCGCTTCCCG TGCCAGGGGT GCTTCTGAAG GAGTTCACTG TSTCCGGCAA 480
CATACTGACT ATCCGACTGA CTGCTGCAGA CCACCGCAA CTGCAGCTCT CCATCAGCTC 540
CTGTCTCCAG CAGCTTTCCC TGTTGATGTG GATCACGCAG TGCTTTCTGC CCGTGTTTTT 600
GGCTCAGCCT CCCTCAGGCG AGAGGCGCTA AGCCACGCTT GGCGCCCTCT CCTAGGTCAT 660
GCCTCCTCCC CTAGGGAATG GTCCAGCAC GAGTGGCCAG TTCATTGTGG GGGCCTGATT 720
GTTTGTCTGT GGAGGAGGAC GGCTTACATG TTTGTTTCTG TAGAAAATAA AACTGAGCTA

```

70 Seq ID NO: 210 Protein sequence:
Protein Accession #: NP_001318.1

75

```

1 11 21 31 41 51
MQAEGRGTTG STGDADGPGG PGIPDGPNGN AGGPGEAGAT GGRGPRGAGA ARASGPGGGA 60
PRGPHGGAAS GLNGCCRCGA RGPESRLLEF YLAMPFATPM EAEIARRSLA QDAPPLEVPG 120
VLLKEFTVSG NILTIRLTAA DHRQLQLSIS SCLQLSLLM WITQCFLPVF LAQPPSQQR

```

80 Seq ID NO: 211 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 52-459

85

```

1 11 21 31 41 51
CCTCGTGGGC CCTGACCTTC TCTCTGAGAG CCGGGCAGAG GCTCCGGAGC CATGCAGGCC 60
GAAGGCGCAG GCACAGGGGG TTCCAGCGGC GATGCTGATG GCCAGGAGG CCCTGGCATT 120
CCTGATGGCC CAGGGGGCAA TGCTGGCGGC CCAGGAGAGG CGGGTGCCAC GGGCGGCAGA 180
GGTCCCCGGG GCGCAGGGGC AGCAAGGGCC TCGGGGCGGA GAGGAGGCGC CCGCGGGGT 240
CGCATGGGC GTGCCGCTTC TGCGCAGGAT GGAAGGTGCC CCTGCGGGGC CAGGAGGCGC 300

```

5 GACAGCCGCC TGCTTCAGTT CCGACTGACT GCTGCAGACC ACOGCCAACT GCAGCTCTCC 360
ATCAGCTCCT GTCTCCAGCA GCTTCCCTG TTGATGTGGA TCACGCAGTG CTTTCTGCC 420
GTGTTTTTGG CTCCAGCTCC CTCCAGGCAG AGGCGCTAAG CCCAGCCTGG CGCCCTTCC 480
TAGGTCATGC CTCTCCCTCT AGGGAATGGT CCCAGCAGCA GTGGCCAGTT CATTGTGGGG 540
GCCTGATTGT TTGTCGCTGG AGGAGGACGG CTTACATGTT TGTTCCTGTA GAAAAATAAG 600
CTGAGCTA

Seq ID NO: 212 Protein sequence:
Protein Accession #: Eos sequence

10 1 11 21 31 41 51
MQAEGQGTGG STGDADGPGG PGIPDGPQGN AGGPGEAGAT GGRGPRGAGA ARASGPRGGA 60
PRGPHGGAAS AQDGRCPGCA RRPDSRLIQF RLTAADHRQL QLSISSCLQQ LSLMMWITQC 120
15 PLPVFLAQAP SGRR

20 Seq ID NO: 213 DNA sequence:
Nucleic Acid Accession #: NM_000555
Coding sequence: 416..1498

25 1 11 21 31 41 51
CTTATTTTAT ATGAATGTGC GATAGCTGCA CCAGCTTGGT GGGGAAAGGG TTTGATGAAT 60
AGCACAAAGA CACTGGCTGT TCCCTGGAGG CTGTCCCTTT AAAGGAGAAT CTTAGTTTAT 120
TCTGGGGGGA GGGGATGCAC ACATTAGAGT AGGAAAGAGG GCTTGAATAA AAATGAAAC 180
ACTCCCCCTT CATAGTCACT GTACTGAAAT GCAGAGACTG CTTCTAAGC TGGAGATGCT 240
AACCTTGGGT AGCTCCTCTT GTTCTCTTCA AGGGGAATTT TGTGAGGCTA TGGATTCAAT 300
TACAACTGTT AGTCATGTGG GCATGTGTGA GGAACAGATG GCCAGTTTAT ATGATTTTAG 360
CCCGAAGTTC CAATTTGATA GGAGCCACTG TCAGTCTCTG AGGTTCCACC AAAATATGGA 420
ACCTGATTTC GGACACTTTG ACGAAGAGAG TAAGACATCC AGGAACATGC GAGGCTCCCG 480
GATGAATGGG TTGCTGAGCC CCACTCAGAG CGCCCACTGT AGCTTCTACC GAACAGAAC 540
CTTGACGGCA TGGTCTGTCA AGAAGAAAGC CAAGAAGGTA CGTTTCTACC GCAATGGGGA 600
CGCTACTTTC AAGGGGATTG TGTACGCTGT GTCTCTGAC CGTTTCTGCA GCTTTGACGC 660
CTTGCTGGCT GACCTGACGC GATCTCTGTC TGACAACTC AACCTGCCTC AGGGAGTGGC 720
TTACATTTAC ACCATTGATG GATCCAGGAA GATCCGAAGC ATGGATGAAC TGGAGGAAGG 780
GGAAAGCTAT GTCTGTTCTT CAGACAACTT CTTTAAAGAG GTGGAGTACA CCAAGAAATG 840
CAATCCCAAC TGGTCTGTCA ACGTAAAGAC ATCTGCCAAT ATGAAAGCCC CCCAGTCTCT 900
GGCTAGCAGC AACAGTGACG AGGCCAGGGA GAACAAGGAC TTGTGCGGCC CCAAGTGGT 960
TACCATCATC CGCAGTGGGG TGAAGCCTCG GAAGGCTGTG CGTGTGCTTC TGAACAAGAA 1020
GACAGCCACG TCTTTTGAGC AAGTCTCTAC TGATATCACA GAAGCCATCA AACTGGAGAC 1080
CGGGGTGTCT AAAAACTCTT ACACCTGCGA TGGAAAACAG GTAACCTGTC TCCATGATT 1140
CTTTGGTGAAT GATGATGTGT TTATTGCTGT TGGTCTGAA AAATTTGCTT ATGCTCAGGA 1200
TGATTTTCTT CTGGATGAAT ATGAATGCGG AGTCATGAAG GGAACCCATC CAGCCACAGC 1260
TGGCCCAAGG GCATCCCAAA CACCTCAGAA GACTTCAGCC AAGAGCCCTG GTCTATGCG 1320
CGAAGCAAGG TCTCCAGCTG ACTCAGCAAA CGGAACCTCC AGCAGCCAGC TCTCTACCCC 1380
CAAGTCTAAG CAGTCTCCCA TCTCTACGCC CACCACTCTT GGCAGCCTCC GGAAGCACA 1440
GGACCTGATC CTGCTCTGAT CCTTGGATGA CTCGCACTCG CTGGTGAATT CCATGTAAAG 1500
GAGGGGAGAG TGCTCAGAGT CCAGAGTACA AATCCAAGCC TATCATGTGA GTAGGGTACT 1560
TCTGCTCAAG TGTCCAACAG GGCTATTGGT GCTTCAAGT TTTTATTTTG TTGTGTGTGT 1620
TATTTTGAAT AACACATTGT AATATGTTGG GTTTATTTTC CTGTGATTTT TCTCTGGGC 1680
CACTGATCCA CAGTTACCAA TTATGAGAGA TAGATTGATA ACCATCTCTT GGGGCAGCAT 1740
TCCAGGGATG CAATAATGTC TAGTCCATGA CCTTTCAATG GAAAGCTTAG GGGCTGGGG 1800
TAAATTTGCC CGTTTAAAT TTGCCCAAC AGTTTCTCTT TTGTAGAGGG GTGTTTAAAT 1860
ATACAGCAAT TAAAGATTTT GTGTGGGGA AAAAAAAT CATTGGCAGA TCCAGAAATG 1920
ACAAACACAA GTGCCCTTTT TCTCTGATC TCAAGAAATG TGGAGGCCCC TGAAGGACA 1980
GCAAGGCAGC TCCCGACGCT CACTCTTCACT TCTGATTGA GGCCGGGGTT TGTGTCCAG 2040
CACCATAATC GGCTGTCAAT GGGGAGAAAT AAACCAACAA CTTATAATG TGACACCAGA 2100
TGCTTAGGAT CCTGTGCTGT GGTAGCTAAG GAGAATAGAC AGAATTGGAA AATAGTCAAG 2160
ACATTTCCGA AGAGTTTATA AAGCACAGTG AATCTCTGGT CAATCTCTCC ACTGAGGCAA 2220
TTTGGAAATCA ATAAGCAATT GATAATAGTT TGGAGTAAGG GACTTCATAT ACCTGATTCC 2280
TCTAGAAGGC TGTCTAATCT ACCACATGAT TACATGAAT GTATGGTATC CATCTATCTC 2340
TGTTCTATTG AATGCCTTGT TAACAGCCAA CACTGAAAC ACTGTGAGAA TTTGTTTTCA 2400
GGCTGACAC CTTTCAGTCT CTTTATATAG CAAGAAATCA ATATCCTTTT TATAAAAAAT 2460
CATGTCTGTA TTTCAGGAGC AAATCTTCA GGCTCCTTTT TTATAAATG GTGATTTTTT 2520
TTTGTCTAA AAAACACATG AAGAAATTT ACCAGAAAAA AAAAAAAG CCGAAGATA 2580
ATGTTATTTA GAAATATGCG TGCTACTGCC AAACAGTAAC CTCCAGGAGA AAACAAGATG 2640
AATAGCAGAG GCCAATTCAA TAGAATCAGT TTTTGTATAG CTTTATAACA GTTATGCTTG 2700
CATTAATAAT TTCAATGTGG ACCAGACATT CTAATTATAT TTTAAATGAA ATGTTACAGC 2760
ATAATTTAAG CAATCTTTT TATCTATAAT CTAATATTT CATACTGAAG ACACAGAAAT 2820
CTTCACTTG TCTTAAACAT TAGAAAGGAT TTCTCTTTAC TAAGACTGA TCAATTGAAA 2880
TAGTTTTCAG TCTTTTGAGA TACAGGTTTA TAACACTGCT TTTTTCCTCC TGTAAACATA 2940
GCCCATATAG GCAAAAACAA CTAATTTTAA TTGAAGGTCT TGCTTGCCAN TCCTGTGTTG 3000
GCTTTNACCA AATAATAAAA TTCCCTTATT CCTTGGTAAT GGTGCAAAAT TTTGAAAGG 3060
CACAGCATCC AAACCAAGCT GCTGTTTGGC TACTGAATGG CTGCAAGTG TTCTCCACT 3120
CTAATGGAAT TGAGCTTGCT GTGTGTGTGT GTGGTGGTGG TGGGAGGGGG TGGTGCATGT 3180
GTGTGTGTGT GTGTGCATCT GCAGCTGCTT CAAATTAAG AAATACTACA AGACACCCCT 3240
GTAATGGAAT GGTGGCACT GGTGGCACT GCTGATGTGC ACTGTGTAGG GGGGAACCCA 3300
GTGTGTGTGG GGTATCTCAA ATGCCCTAG ACAAGCTTCA GATGTCTGTA GCTACCAAAA 3360
ACATTTTCGG TTCAAGAAAA GTGAGATGAT GGTAGTACTG GTTCTGGT AAATTGAAAA 3420
ACCCCAATG ATGAGGATCT CTTTGTGCC CTCTCCTTT TTTTGTAAAC CCAATCAAAA 3480
CCATTAATAA GCCCAATTTA CTAANCCCT ATTTCTTTCT AGAAGCTCAG GGTTTNCTTA 3540
GTGCTCCCA NAACATTTG TAGTTAATTG GGAAGAAAGT ATACTTGGAT TAGGGGGTGT 3600
GGCATAAAG AATGGTGGGA GGCCTGATT TAAATTCAG GCGAAGACC CCAATGACTC 3660
CACCCATAGT NTCACCTTAG GTCTCATTTA GTCCATCACC TTTATTTTAA GTTGAGGAAG 3720
TGGAGGCTGG TAAAGAGCAG GACCAGAGGA AGAATCCAGA TTTCTTATG CTTGGGCTC 3780
ACACTAGCTC TGTGATATT TCCTTGATTG CGGTATATGT ACTACTAGAA AATACCAAT 3840
85 GGATATATT TCTTTAGGAT AACCTTTGAA CCAACAATNT TCAATAACA TAGTACATCT 3900

	TCATCTTAC	TTTTAATCGA	GTATAAGGAA	ATGTTTCTTT	ATGGCCATT	TGGAGGGAGC	3960
	AGGGGATGAG	GCTTGGCATA	GTCCAAAATT	TAAGNCTCCA	ATAATTAAAT	GCAITTTAAA	4020
	TTGTTTTAAA	TTGGCCCACT	TTCAAGGCAA	TTTTTTTTGT	GTGTCGTAA	CTGAGCTCCT	4080
5	CCACCCCTGT	CATTCACTTC	CAATTTTACC	CAATCCAATT	TTAGCACTCA	AGTTCCTATG	4140
	TGTTAATTTT	TGCAAGGTCT	ACACACATCA	AGTCAGCAAG	CATTGGCCAC	CACCTCCCTAT	4200
	ACTTCTCCCT	CTTTTGTACA	CACACACACA	CACACACACA	CACAAATCCAT	CTCTTGCTTG	4260
	TTCTTACCTC	CCTGATTTTT	CTTCCCTACA	GAAATAGAAA	TAGGGACAAA	GAAGGGGAAA	4320
	ATGTATATAT	TGGGCTGGG	CTGAACAAC	AACTTCATAA	GTAGTATTAA	CTAGGGGTAA	4380
10	ATTGAGAGAA	AAGCTCCTTT	TCTCTTCACT	GTTTTGGAAA	GGATAGCCAT	TAGCATGACT	4440
	GCTTTGTGTC	CTATGGAACT	TTAGTATTAG	CCTAGATTGA	ATTATAGCGT	TTTTCTAGCT	4500
	GAAGGAACCT	TAAGATCACA	TCATCTACTC	CTCTACTCCA	AATTTCTCAT	TCTTCAGGCC	4560
	AGGAAACCGA	GACACAGAGG	TAAAGTAATT	TCCCAAGGT	CACACAGCTG	GCTGGGGCAG	4620
	GATTGGGTTT	ACAAACCCACA	TCTCTGGCT	CTTATTCAG	GGCCTTTTCC	CACTAAGTAG	4680
	TATTGCTTC	CATTAGGCTC	CTGAGAGTTA	TTTCTCAGG	TCATGTGCA	TCTTGGAGCC	4740
15	ACATGCTGTC	GCCTTGATCT	CAGTGGGAAA	TNCAACCCAGC	AACCTAATAC	AGCCCTTTT	4800
	CCCTGCAATC	ACCTGGTTC	CATCCACATG	GGTTCAGAT	GTCTTGAAG	AGAGTAGAGC	4860
	ATTGAGGCC	AATAGGAGCA	ATGGGGTCC	TGGCCTTGTC	CATCTGATTC	AGGAGATCAC	4920
	TGCTCCATCG	TGAGSAGGCC	TCTGAATAGC	CCCCACTGA	ATGCTTGCT	TGCCCAATG	4980
20	GAATGGAGGA	AGATTGATTT	TCTCCATCAG	TTCCACTTGT	GTCACTCAT	AATGGTTGGT	5040
	CTTTCCAGGC	TGAGGGAAT	GTTTCTTGT	TCCANAGTAN	AAAAAAGAAA	GAGTGGAAACA	5100
	ATANTCTTGT	CTACTCTAAC	TTTCTGAGAT	GGCTTTTCAA	CATTTAAAAA	AAACTAGTGT	5160
	GGTACCATTG	ACTGGCANGA	TTNTTTTAG	AATATGGGAG	TAAGATGAGG	TAGAGAAAAT	5220
	AACCTGGTCT	CACGTGGTGT	GCCCTCATCC	ACAATGTCCC	CAAGCCATC	CTGCTNIGAT	5280
25	GAGGACAAT	TCCAGGTATA	AGCAAGGGGC	TTTGTGACAA	AAATGTACCC	TGGCTGATGT	5340
	TAAACATGTC	CTCCTGTGTT	TGCACCAAAA	TAGCAAGCTG	TGTGCTCTAT	ACACTCTTCC	5400
	CATGCTCTTG	TGTACACTGC	TCTGTGGGCC	TTCCACAGCA	GAACCCAGGG	CAAAAGGGTC	5460
	CAACACATG	GTTTCTCTTG	CTGCAAGGCT	NTTCTGGGA	ACTAAGGGGG	TATTTATTAG	5520
	TTCACTTNTA	AGAGACCTCT	TTCTGGGCTT	ACCCCACTCC	TCAGTACTT	CTCTCTCCTT	5580
30	CCTCTCTCTC	CTCCACAGCT	ACAAGTAACC	AAGGAACCTG	AAAGTGGATG	TGTAGCTATT	5640
	TGAAGAAGGC	AAGGAACCTT	GAGATTCTTC	TTTGAATCCT	TTAGTCCAA	TCTTAGACCA	5700
	GTGATTGGTG	CTTACCTTGA	ACAAAATTTT	GTCTGTGTTT	CTAATCCCTT	CAATACINTG	5760
	GGTACAATGC	TCCCAATCAC	CCTGCACATT	TGATTCTAAA	TGGCTTTTAT	TTTTTAAAAA	5820
	TCCATATCCC	TAGGACAGAA	NAACAGGATG	CCTATATCCC	CAAAATGAGC	TCCAGSACAC	5880
35	TGATGGGAAT	GATCCCAANG	ATCACCACAC	CTCAGAAAAC	GTCTGTGCCA	ANAGACTTCC	5940
	CCAGATAGAA	NACACTGGGAC	AGTGGTTTGA	ACGACTTCTT	TTATGGTTGT	CCAGTTTGCT	6000
	ATGGAAATAA	AAGGCATTTA	TTTTTAAAAA	AAGATGATTT	GAACCTGTCT	TGCGCCACAT	6060
	AGGGCCACTT	GGATCCATTT	CCAGGCCCTA	CTCATATATT	GCCTTCACTG	AAGGCGTTTG	6120
	GCTTTAAGTC	CCAGACTGGT	CTCCCAAGTG	AACCATAGT	GTTTTGGAGC	TCATCTGGGG	6180
40	TGAGGCATGA	GAATGTTGCC	CCATCTATCC	CTTCAGGAAA	AGGTGCTTTC	CCTCCCTTTC	6240
	TCCTAAAGCC	TGGTCCCCAA	AAATGTTTTT	TGCTCCAAA	AGTCTAGTAT	GGTCTTTATA	6300
	CACCCANACT	CTTAGTGTG	CGTCTGCTCT	TGTTTCTCTG	TTAAGGATCT	ATGCANACT	6360
	CCCGCTTTGG	CTTAGCTAGC	GTGACATTGG	CTATCAITTG	ACAAGACTAA	CTTTTTTTTT	6420
	TTTTTTTTTG	ACTGAGTCTC	CCTCTGTAC	CTAGGCTGGA	GTGCAAGTGC	ACAATCTTGG	6480
45	CTCGCTGCAA	CCTTCAACCT	TCACTCCCA	GGTGAAGCG	ATTCTCTGTC	CTCAGTCTCC	6540
	CGAGTAGCTG	NCATTACAGG	CGTGCACCAC	CAAACTGGC	TATTTTITTA	TTATTATTAT	6600
	TTTTAGTAGA	GATGGGGTTT	CACCATGTTG	GCCAGACTGG	TCTTGAATCT	TGCGCTCAA	6660
	ATTATCTGCC	CACCTCGGCC	TCCCAAGTG	CTGGGATTAC	AGGCATGAGC	ACCATGCCCA	6720
	GCTGACAGAA	CTAATTTTTT	ATCCCTTGGT	TTATGGCTT	CAACATCTTC	TGGAATCAGA	6780
50	GGTGATTTTT	TCTTACCTTG	GATGCTTGAG	ACTAGGGGAG	TATAGAATTC	CAATGGGTAA	6840
	TTAAGGCATC	TTCTGTCTCC	TGATCAGAA	GGCAGTTTAT	TTGGGAGAGG	TCAGATGGCA	6900
	CAACAGAACT	CACCTGTAA	GTAAGGCAAA	GACTTTGAAG	GCATTAGCGT	TTCTCATTAC	6960
	TTAGTCAAT	AACCTGTAGG	GAATCAATGG	CTTTTTTGGC	GCTCTAOCCT	TTTGTGTATC	7020
	TCTTTGACTT	TTCTTCTCTC	GTCTAGTTTC	CTCTGTTCTC	AGTTTATATT	CTATGTTATC	7080
55	AGTCTCTCTT	TCCACAGTAC	AAACATCCAT	CCTTCTCTCT	GTGCAATCTT	GTCTCTCCCT	7140
	CTTATTATCT	TTATTGTATC	TTTTTCTTTC	CTCCCTGTCT	AGGCATTGGG	CATGTGCTCT	7200
	TTCTTAGCCT	GTGATTTTTC	CTGGGACTG	ATGATAAATT	ATTTCAGAT	TCAATCAGCC	7260
	CTGCTCTTAC	CCAGTCCCAA	TCAGAAATAT	GTGGTGGGG	AATCAACCTG	ATCCTGGCCC	7320
	TTCTCTCTTC	TCCATTTTCA	TTCTGAATCC	CCCTCAGCAG	ATCTTTACAA	GCAGTTTCTT	7380
60	TATAGCTCAT	GTATCTTTAG	CTTCTTGCTT	TCCAGCACT	GTACAGAAAT	CTTTGTGGTT	7440
	CCTTTTGTAG	CTGACATTTT	GTGGAGCAGT	GAAGCGTGCT	CAGAGACATA	ATCAGCTGAA	7500
	GAGAAAAAAT	CCACCCATGG	ATTATATATCA	GCTAAATACT	AATAATTGAT	TTGTTTGTAT	7560
	GTGCCCCATA	TTTAAAGC	TGCAATATAA	TATAATGAGG	GACCACAGGT	AATTTCTCCT	7620
	GTCAATTTGTT	TTGGCTGGAT	GGGGGTGGGG	GAGTAATTGC	TTAAAGTTT	ACCATTACAC	7680
65	ATTAACCTCT	CTATAATAAT	CTTGTTTGGG	GCTTGCTAAC	TGTTGAGCTG	TTTTAACTAA	7740
	ACTGGTAGGC	AATCGGAGTT	GATTFAAATG	AAAAGATAAT	TTAACAAATC	TATACTATAA	7800
	AAAGAGACAT	TGCTTAAAT	GACATGTATT	TTTTCTTCT	GAGTCACTTA	AACATTACT	7860
	CTTGACACCA	ACTGTTCTAT	ATACTGAATA	GACAGTCCAT	ATAAGAGAAA	TTAGTGGACC	7920
	TAAAGAAGCC	AGATTGTAGG	TGTTAATTTA	TTAAACAGAA	TTGCAAGGCC	CTTGGAAATG	7980
70	TCATGTCTTG	GCAATACCAT	ATGGCATGCC	AAAATTTACA	ATGACTTTTC	TTTATAAGTT	8040
	ATCCAAAAGG	GATTGAAACA	AGTAAGAGGT	TATGCCAAAA	TGCTCTCAAT	GTATGGTCTT	8100
	GTAATATATT	GCAGCTTGAA	GCCAATGATC	CCTTATGACT	TGTATACAA	TAATGCAATG	8160
	TTTATGTAAT	TTTGCAITTT	CCAAGTGTGG	TAACTCTTTA	AAATGTTTTT	GATCAACCTT	8220
	MTGTGCCATT	AAACTTGTAC	AGAAAATGTT	TTTATGGCCA	TTTTCAAAGG	GAGAAAGTTT	8280
75	AAAATGGAAA	CAGCCCAACC	TTTCTGCCCT	ATAGCTGTAG	TTAGAATTGA	GTACCTGTAG	8340
	CAAAACAGCT	GTAATTTGGT	GTTGTAGTGT	TAGAGGTGTT	AGCTTGTCTAG	TGACTAGCTT	8400
	TGGAGAGTAA	ATGCATGGTA	TTGTACATCA	CATTCTTTAA	CTCGTTTTAA	CCTCTGAAAA	8460
	GAATATATTC	TTCTTTGTAG	TCTTCTTCC	CACCCCTTGG	CCCTCTCCCT	CTCCCTGCTC	8520
	CCAGTTGTCT	TACAGTTGTA	AATATCTGAT	TTGAGGCCCA	ATAACTCTTG	CCAAGTAAAG	8580
80	TCAGCAACA	ACAAACAACA	CAAAATGTGG	GGAAAAGGCA	TTTCTCAACC	ATCTCTCAGC	8640
	AGTTATTGAT	CATTTCTTAA	GGAAACAGCAT	TGTGATCAAA	GACTCAACTT	TACGTAAAAA	8700
	TCAGTGTGTA	ATTGGGGTGG	TATTGGCCAT	TGATTACATT	CAGGATTGAA	TAGTTTTCAG	8760
	AATCACATGT	AATCCAAAGA	CAGTAGGTAG	TGATGTCCT	TATCCCTGCA	GCTGTTTTAA	8820
	GATAGAGACC	TCAGAAGACT	CTGCTTGACC	GATGACCAAT	AATTATTGTA	AAAAAAGA	8880
85	AAAAATGAGA	GAATAAAAAC	AGATATTATA	GAACCTTAGC	CACCTATTTA	GAATAGTTAT	8940
	AGCCAGAAA	AAAAACAGG	GCAATGATTC	AAATGCATTA	CTATCAGTGT	CCTAGGCAAT	9000
	ACCTAACCTA	CTCTGAAATT	GTGATTCAAA	AGCAGTATTT	CAAGAGGCAT	TCTCCTTTTT	9060
	TGGTTTGGCT	ACCCCACTTG	GACTGGTAGG	TTTGGTGAGG	CCCCATAAA	CCAGCTGGAG	9120

WO 02/086443

PCT/US02/12476

CAGACCCCTTT TCATCTCCTG TGCCGTGAAC ACCCCTCTTC CCCCACCCCC TCCGCAATTC 9180
AATGAGGGCT TTCTTGGGTC AGAGGACTTC AAGGTTGTCT AGAGAAGTTT GCCATGTGTG 9240
TAAGGTGCTG TGAACCTGTA GTGCTGAAGA TTGCGAGCAT TCAATACCAG GCAGCCAAAAG 9300
AGCTGCTCTT GCAATTAATT TGGCTCTCAA GCTCTGTTCT TCATCGCATT CTCATTCTCG 9360
TGTACATTG CAAGATGTGT GTAATGTCAT TTTCCAAAA TAAATTTGA TTTCAAT

Seq ID NO: 214 Protein sequence:
Protein Accession #: NP_000546

1 11 21 31 41 51
MELDFGHFDE RDKTSRNMRG SRMNGLPSP T HSAHCSFYRT RTLQALSNEK KAKKVRFYRN 60
GDRYFKGIVY AVSSDRFRSF DALLADLTRS LSDNINLPQG VRYIYITDGS RKIGSMDELE 120
EGESYVCSSD NEFKKVEYTK NVNPNWSVNV KTSANMKAPQ SLASSNSAQA RENKDFVRPK 180
LVTIIRSGVK PRKAVRVLLN KKTAKHFEQV LTDITEAIKL ETGVVKLYT LDGKQVTLH 240
DFFGDDVFI ACGPEKFRYA QDDFSLDENE CRVMKGNPSA TAGPKASPTP QKTSKSPGP 300
MRRSKSPADS ANGTSSSQLS TPKSKQSPIS TPTSPGSLRK HKDLVLP LSL DDSLSLGDSM

Seq ID NO: 215 DNA sequence
Nucleic Acid Accession #: NM_130467
Coding sequence: 312..644

1 11 21 31 41 51
GGCAGGAGGC AGAGCTCTGC AAGGAGAGGT TGTGTCTTCG TTCCTTCCGC CATCTTCGTT 60
CTTTCCAAAC TCTTCGTTCT TTCTCACTGA CCGAGACTCA GCGGTAGGT CTGCAGAGTG 120
GTCTTCTCTG TAATTTAGTT GTGAGTGAAT GTGTGGAGGA GCCAGCGGGC TTAGACAGG 180
TCCTGTGGCA CAGTCCGTGG CTTTGAGGGA AAAGGGCCTC GGGTGTGTCC TCCGCTTCC 240
CCCAGGTCGT GATGACGGCG CCATGGGCGG GTAATCGTGG CTGGGCTGGA ACGAGGGAGG 300
AAGTGAGAGA TATGAGTGAG CATGTAACAA GATCCCAATC CTCAGAAAGA GGAATGACC 360
AAGAGTCTTC CCAGCCAGTT GGACCTGTGA TTGTCCAGCA GCCCACTGAG GAAACACGTC 420
AAGAAGAGGA ACCACCAACT GATAATCAGG GTATTGCACC TAGTGGGGAG ATCAAAAATG 480
AAGAGCACC TGCTGTTCAA GGGACTGATG TGAAGGCTTT TCAACAGGAA CTGGCTCTGC 540
TTAAGATAGA GGATGCACCT GGAGATGGTC CTGATGTCAG GGAGGGGACT CTGCCCACTT 600
TTGATCCAC TAAAGTCTGT GAAGCAGGTG AAGGGCAACT ATAGGTTTAA ACCAAGACAA 660
ATGAAGACTG AAACCAAGAA TATGTCTCTT ATGCTGGAAT TTTGACTGCT AACATTCTCT 720
TAATAAGATT TTACAGTTTT CTGCAAAAAA AAAAAAAAAA AAA

Seq ID NO: 216 Protein sequence:
Protein Accession #: NP_569734

1 11 21 31 41 51
MSEHVRISQS SERGNDQESS QVVGPFVIVQ PTEEKREBEE PPTDNQGLAP SGEIKNEGAP 60
AVQGTDEAF QQELALLKIE DAPGDGPDVR BGTLPFTDPT KVLEAGEGQL

Seq ID NO: 217 DNA sequence
Nucleic Acid Accession #: NM_001476.1
Coding sequence: 82..435

1 11 21 31 41 51
GCCAGGAGGC TGTGAGGAGC TGCTGTGTGG TTCCTGCCGT CCGGACTCTT TTCCTCTAC 60
TGAGATTTCAT CTGTGTGAAA TATGAGTTGG CGAGGAAGAT CGACCTATTA TTGGCCTAGA 120
CCAAGGCGCT ATGTACAGCC TCCTGAAGTG ATTGGGCCTA TGCGGCCCGA GCAGTTCACT 180
GATGAAGTGG AACCAGCAAC ACCTGAAGAA GGGGAACCGA CAACTCAAGT TCAGGATCCT 240
GCAGCTGCTC AGGAGGGAGA GGATGAGGGA GCATCTGCAG GTCAAGGGCC GAAGCCTGAA 300
GCTGATAGCC AGGAACAGG TCACCCACAG ACTGGGTGTG AGTGTGAAGA TGGTCTCTAT 360
GGGAGGAGG TGGACCGGCC AAATCCAGAG GAGGTGAAAA CGCCTGAAGA AGGTGAAAAG 420
CAATCAGAT GTTAAAGAA GACACGTTGA AATGATGCAG GCTGCTCTTA TGTGGAAT 480
TTGTTCATTA AAATCTCTCC AATAAGCTT TACAGCCTTC TGCAAAA

Seq ID NO: 218 Protein sequence:
Protein Accession #: NP_001467.1

1 11 21 31 41 51
MSWRGRSTYY WPRPRRYVOP PEVIGPMRPE QFSDEVEPAT PEEGEPATOR QDPAAAQEGE 60
DEGASAGQGP KPEADSQEQG HPQTGCECED GPDGQEVDP NPEEVKTPEE GEKQSQC

Seq ID NO: 219 DNA sequence
Nucleic Acid Accession #: NM_001476
Coding sequence: 90..3671

1 11 21 31 41 51
ACAGCGGAGC GCAGAGTGAG AACCAACAAC CGAGGCGCGG GGCAGCGACC CCTGCAGCGG 60
AGACAGAGAC TGAGCGGCCC GGCACGCCCA TGCTGCGCT CTGGCTGGGC TGCTGCCTCT 120
GCTTCTCGCT CCTCTGCCC GCAGCCCGGG CCACCTCCAG GAGGAAAGTC TGTGATTGCA 180
ATGGGAAGTC CAGGCAGTGT ATCTTTGATC GGGAACTTCA CAGACAAACT GGTAAATGGAT 240
TCCGCTGCTT CACTGCAAT GACAACACTG ATGGCATTCA CTGCGAGAAG TGCAAGAAATG 300
GCITTTACCG GCACAGAGAA AGGAGCCGCT GTTTGCCCTG CAATTGTAAC TCCAAAGGTT 360

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

CTCTTAGTGC TCGATGTGAC AACTCTGGAC GGTGCAGCTG TAAACCAGGT GTGACAGGAG 420
CCAGATCGCA CCGATGTCTG CCAGGCTTCC ACATGCTCAC GGATGCGGGG TGCACCCAAG 480
ACCAAGAGACT GCTAGACTCC AAGTGTGACT GTGACCCAGC TGGCATGCSA GGGCCCTGTG 540
ACGCGGGCCG CTGTGTCTGC AAGCCAGCTG TTACTGGAGA ACCTGTGTAT AGGTGTGTAT 600
CAGGTTACTA TAATCTGGAT GGGGGGAACC CTGAGGGCTG TACCCAGTGT TTCTGTCTATG 660
GGCAATTCAGC CAGCTGCGCG AGCTCTGCAG AATACAGTGT CCATAAGATC ACCTCTACTCT 720
TTCATCAAGA TGTGTGATGG TCGAAGGCTG TCCAAAGAAA TGGGTCTCCT GCAAAGCTCC 780
AATGGTCACA GCGCCATCAA GATGTGTTTA GCTCAGCCCA ACGACTAGAC CCTGTCTATT 840
TTGTGGCTCC TGCCAAATTT CTGGGGAATC AACAGGTGAG CTATGGGCAA AGGCTGTCTCT 900
TTGACTACCG TGTGGACAGA GGAGGCAGAC ACCCATCTGC CCATGATGTG ATCTGTGAAG 960
GTGCTGGTCT ACGGATCACA GCTCCCTTGA TGCCACTTGG CAAGACACTG CCTTGTGGGG 1020
TCACCAAGAC TTACACATTC AGGTTAAATG AGCATCCAAG CAATAATTGG AGCCCCCAGC 1080
TGAGTTACTT TGAGTATCGA AGGTTACTGC GGAATCTCAC AGCCCTCCGC ATCCGAGCTA 1140
CATATGGAGA ATACAGTACT GGTGTACTTG ACAATGTGAC CCTGATTCA GCGCCGCCCTG 1200
TCTCTGGAGC CCAGACACCC TGGGTGAAAC AGTGTATATG TCCTGTGGGG TACAAGGGGC 1260
AATTCTGCCA GSAATGTGCT TCTGGCTACA AGAGAGATTC AGCGAGACTG GGGCCCTTTG 1320
GCACCTGTAT TCCTTGTAAAC TGTCAAGGGG GAGGGGCTGT TGATCCAGAC ACAGGAGATT 1380
GTTATTCAGG GGATGAGAAT CCTGACATTG AGTGTGCTGA CTGCCCATT GGTTTCTACA 1440
ACGATCCGCA GAGCCGCCGC AGCTGCAAGC CATGTCCCTG TCATAAGCGG TTCAGCTGCT 1500
CAGTGTAGCG GAGAGACGAG GAGGTGGTGT GCAATAACTG CCTCCCGGG GTACCCGGTG 1560
CCCGCTGTGA GCTCTGTGCT GATGGCTACT TTGGGGACCC CTTTGGTGAA CATGGCCAG 1620
TGAGGCTTGT TCAGCCCTGT CAATGCAACA ACAATGTGGA CCCAGTGCC TCTGGGAATT 1680
GTGACCGGCT GACAGGCAGG TGTGTGAAGT GTATCCACA CACAGCCGCG ATCTACTGCG 1740
ACCAAGTCAA AGCAGGCTAC TTGCGGGACC CATTTGGCTCC CAACCCAGCA GACAAGTGTG 1800
GAGCTTGCAG CTGTAAACCC ATGGGCTCAG AGCCTGTAGG ATGTGGAAGT GATGGCACCT 1860
GTGTTTGCAA GCGAGGATTT GGTGGCCCCA ACTGTGAGCA TGGAGCATT AGCTGTCCAG 1920
CTTGCTATAA TCAAGTGAAG ATTGAGATGG ATCAGTTTAT GCAGCAGCTT CAGAGAAATG 1980
AGGCCCTGAT TTCAAAGGCT CAGGGTGGTG ATGGAGTAGT ACCTGATACA GAGCTGGAAG 2040
GCAGGATGCA GAGGCTCTAG CAGGCCCCCTT AGGACATTCT GAGAGATGCC CAGATTTTCA 2100
AAGTGTCTAG CAGATCCCTT GGTCTCCAGT TGGCCAAAGT GAGGAGCCAA GAGAACAGCT 2160
ACCAAGGCCG CCTGGATGAC CTCAGATGA CTGTGGAAAG AGTTCGGGCT CTGGGAAGTC 2220
AGTACCAGAA CCGAGTTCGG GATACTCACA GGCTCATCAC TCAGATGCAG CTGAGCCTGG 2280
CAGAAAGTGA AGCTTCTCTG GGAAACACTA ACATTCTCTG CTCAGACCAC TACGTGGGGC 2340
CAAATGGCTT TAAAGTCTGT GCTCAGGAGG CCACAAGATT AGCAGAAAGC CACGTTGAGT 2400
CAGCCAGTAA GAGACTGCAA CTGACAAGGG AAACCTGAGGA CTATTCCAAA CAAGCCCTCT 2460
CACTGTGCGC CAAGGCCCTG CATGAAGGAG TCGGAAGCGG AAGCGGTAGC CCGGACGGTG 2520
CTGTGGTGCA AGGGCTTGTG GAAAAATTGG AGAAAACCAA GTCCCTGGCC CAGCAGTTGA 2580
CAAGGGAGGC CACTCAAGCG GAAATTGAAG CAGATAGGTC TTATCAGCAC AGTCTCCGCC 2640
TCCTGGATTG AGTGTCTCGG CTTGAGGAG TCAGTGTATG GTCTTTTCA GTGGAAGAAG 2700
CAAGAGGAT CAACAAAAA GCGGATTAC TCTCAACGCT GGTAAACAGG CATATGGATG 2760
AGTTCAGCG TACAAAAAG AATCTGGGAA ACTGAAAGA AGAAGCACAG CAGCTCTTAC 2820
AGAATGAAA AAGTGGGAGA GAGAAATCAG ATCAGCTGCT TTCCCGTGCC AATCTTGCTA 2880
AAGCAGAGC ACAGAAGCA CTGAGTATGG GCAATGCCAC TTTTATGAA GTTGAGAGCA 2940
TCCTAAAAA CCTCAGAGAG TTTGACCTGC AGGTGGACAA CAGAAAGCA GAAGCTGAAG 3000
AAGCCATGAA GAGACTCTCC TACATCAGCC AGAAGGTTTC AGATGCCAGT GACAAGACCC 3060
AGCAAGCAGA AAGAGCCCTG GGGAGCGCTG CTGCTGATGC ACAGAGGGCA AAGAATGGG 3120
CCGGGGAGGC CCTGGAATTC TCCAGTGAGA TTGAACAGGA GATTGGGAGT CTGAACCTGG 3180
AAGCCAAATG GACAGCAGAT GGAGCCTTGG CCAATGAAAA GGGACTGGCC TCTCTGAAGA 3240
GTGAGATGAG GGAAGTGGAA GGAGAGCTGG AAAGGAAGGA GCTGGAGTTT GACACGAATA 3300
TGGATGCAAT ACAGATGGTG ATTACAGAAG CCCAGAAGGT TGATACCAGA GCCAAGAAAG 3360
CTGGGGTTAC AATCCAAGAC AACTCAACA CATTAGACGG CCTCTGCAT CTGATGGAAC 3420
AGCCTCTCAG TGTAGATGAA GAGGGGCTGG TCTTACTGGA GCAGAAGCTT TCCGAGGCCA 3480
AGACCCAGAT CAACAGCCAA CTGGGGCCCA TGATGTGAGA GCTGGAAGAG AGGGCACGTC 3540
AGCAGAGGGG CCACCTCCAT TTGCTGGAGA CAAGCATAGA TGGGATTCTG GCTGATGTGA 3600
AGAACTTGA GAACTTAGG GACCAACCTGC CCCCAGGCTG CTACAATACC CAGGCTCTTG 3660
AGCAACAGTG AAGCTGCCAT AAATATTCT CACTGAGGT TCTTGGGATA CAGATCTCAG 3720
GGCTCGGGAG CCAATGTCATG TGAGTGGGTG GGATGGGGAC ATTTGAACAT GTTTAATGGG 3780
TATGCTCAGG TCAACTGACC TGACCCCATT CCTGATCCCA TGGCCAGGTG GTTGTCTTAT 3840
TGCAACATAC TCCTTGTCTC CTGATGCTGG GCAATGAGGC AGATAGCACT GGGTGTGAGA 3900
ATGATCAAGG ATCTGGACCC CAAAGAATAG ACTGGATGGA AAGACAAACT GCACAGGCAG 3960
ATGTTTGCCT CATAATAGTC GTAAGTGGAG TCCTGGAATT TGGACAAGTG CTGTGGGAT 4020
ATAGTCAACT TATTTCTTGA GTAATGTGAC TAAAGGAAAA AACTTTGACT TTGCCAGGC 4080
ATGAAATTCT TCCTAATGTC AGAACAGAGT GCAACCCAGT CACACTGTGG CCAGTAAAT 4140
ACTATTGCTC CATATTGTCC TCTGCAAGCT TCTTGTGAT CAGAGTTCTT CCTACTTACA 4200
ACCCAGGGTG TGAACATGTT CTCCATTTC AAGCTGGAAG AAGTGAGCAG TGTGGAGTG 4260
AGGACCTGTA AGGCAGGCC ATTACAGAGT ATGGTGTCTG CTGGTGCCTG CCACCTTCAA 4320
GTTCTGGAGC TGGGCATGAC ATCCTTTCTT TTAATGATGC CATGGCAACT TAGAGATTGC 4380
ATTTTATTA AAGCATTTCC TACCAGCAA GCAAAATGTT GGAAGTATT TACTTTTTCG 4440
GTTTCAAAGT GATAGAAAAG TGTGGCTTGG GCATTGAAAG AGGTAAAAAT CTCTAGATT 4500
ATTAGTCTTA ATTCAATCCT ACTTTTCGAA CACCAAAAT GATGCGCATC AATGTATTT 4560
ATCTATTTT CTCAATCTCC TCTCTCTTTC CTCCACCCAT AATAAGAGAA TGTTCTTACT 4620
CACACTTCAG CTGGGTCACA TCCATCCCTC CATTATCCT TCCATCCATC TTTCCATCCA 4680
TTACCTCCAT CCATCTCTCC AACATATATT TATTAGTAC CTACTGTGTG CCAGGGGCTG 4740
GTGGGACAGT GGTGACATAG TCTCTGCCCT CATAGAGTTG ATTGTCTAGT GAGGAAGACA 4800
AGCATTTTAA AAAAATAAAT TTAACCTTAC AAACCTTTGT TGTCACAAGT GGTGTTTATT 4860
GCAATAACCG CTGPGTTTGC AACCTCTTTC CTCACAGAA CATATGTTGC AAGACCCCTC 4920
CATGGGGGCA CTTGAGTTT GGCAGGCTG ACAGAGCTCT GGGTTGTGCA CATTCTTTG 4980
CATTCAGCT GTCACTCTGT GCCTTTCTAC AACTGATTGC AACAGACTGT TGAGTTATGA 5040
TAACACCACT GGGAAATGCT GGAGGAACCA GAGGCATTC CACCTTGGCT GGGGAAGACT 5100
TGTGCTGCC TTGCTCTGT ATTTCTTGG ATTTCTCTGA AAGTGTTTT AAATAAGAA 5160
CAATTGTAG ATGCC

Seq ID NO: 220 Protein sequence:
Protein Accession #: NP_005553

1. 11 21 31 41 51

WO 02/086443

PCT/US02/12476

MPALWLGCCCL CFSLLPAAAR ATSRREVDCD NGKSRQCPD RELHRQTNGG FRCLNCNDNT 60
 DGIHCEKCKN GFYRRERDR CLPCNCKSKG SLSARCDNSG RCSCKPGVTG ARCDRLPGF 120
 HMLTDAGCTQ DQRLDLSKCD CDPAIAGPC DAGRCVCKPA VTGERCDRCR SGYYNLDGGN 180
 PEGCTQCFCY GHSASCRSSA EYSVHKITST FHQDVGWKA VQRNGSPAKL QWSQRHQDVF 240
 SSAQRDPVY FVAPAKFLGN QQVSYGQSL SFDYVRDGRG HPSAHDVILE GAGLRITAPL 300
 MPLGKTLPCG LTKYTTFRLL EHPSSNWSPO LSYFEYRRL RNLTLALRIRA TYGEYSTGYI 360
 DNVTLISARP VSGAPAPWVE QCICPVGYKG QPCQDCASGY KRDSARLGPF GTCIPNCQGG 420
 GGACDPDTGD CYSGDENPDI ECADCPIGFY NPDHDPSCCK PCPCNPGFSC SVMPEEEVV 480
 CNXCPGVTG ARCELCAADY FGDPFGHEGP VRPCQPCQCN MNVDPASGN CDRLTGRCLK 540
 CIHNTAGIYC DQCKAGYFGD PLAPNPAKDC RACNCPMGS EPVGCSDGT CVCKPGFGGP 600
 NCERGAFCSP ACYNQVKIQM DQFMQQLQRM EALISKAQGG DGVVPDTELE GRMQQAEQAL 660
 QDILRDAQIS EGASRSLGLQ LAKVRSQENS YQSRLLDLKM TVERVRALGS QYQNRVRDTH 720
 RLITQMQLSL AESEASLGNT NIPASDHYVG PNGFKSLAQE ATRLAESHVE SASNMEQLTR 780
 ETEDYSKQAL SLVRKALHEG VSGSGSGPDG AVVOGLVEKL ETKSLAQQL TREATQAEIE 840
 ADRSYQHSIR LLDVSVRISG VSDQSFQVEE AKRIKQKADS LSTLVTREND EFKRTQKNGI 900
 NWKEAQQLL QNGKSGREKS DQLLSRANLA KSRAQALSM GNATFYEVES ILKNLREPLD 960
 QVDRKABAE AMKRLSYIS QKVSADSKT QQAERALGSA AADAQRAKNG AGEALRISSE 1020
 IQEIGSLNL EANVTADGAL AMEKGLASLK SEMREVEGEL ERKELEFDTN MDAVQMVITE 1080
 AQKVDTRAKN AGVTIQTDLN TLDGLLHLM QPLSVDEEGL VLLQKLSRA KTGINSQLRP 1140
 MMSELEERAR QQRGHLHLE TSIDGILADV KNLENIRDNL PPGCYNTQAL EQQ

Seq ID NO: 221 DNA sequence
 Nucleic Acid Accession #: NM_016529
 Coding sequence: 13-1854

1 11 21 31 41 51
 GTCAAGAAAA GAATGTCGTG AATTGTCGGA ACTCCTTCAG GACGACTTCG GCTTTACTGT 60
 AAAGGGGCTG ATAATGTGAT TTTTGAGAGA CTTTCAAAAG ACTCAAAATA TATGGAGGAA 120
 ACATTATGCC ATCTGGAATA CTTTGCCACG GAAGGCTTGC GACTCTCTGT TGTGGCTTAT 180
 GCTGATCTCT CTGGAATAA GATATGAGAG TGGCTGAAAG TCTATCAGGA AGCCAGCACC 240
 ATATTGAAAG ACAGAGCTCA ACGGTTGGAA GAGTGTACG AGATCATTGA GAAGAATTG 300
 CTGCTACTTG GAGCCACAGC CATAGAAGAT CGCCTTCAAG CAGGAGTTC AGAAACCATC 360
 GCAACACTGT TGAAGGCAGA AATTAAATA TGGGTGTGA CAGGAGACAA ACAAGAACT 420
 GCGATTAAATA TAGGGTATTG CTGCCGATTG GTATCGCAGA ATATGGCCCT TATCCTATTG 480
 AAGGAGGACT CTTTGGATGC CACAAGGGCA GCCATTACT AGCACTGACG TGACCTTGGG 540
 AATTGCTGGG GCAAGGAAAA TGACGTGGCC CTCTCATCG ATGGCCACAC CCTGAAGTAC 600
 GCGCTCTCCT TCGAAGTCCG GAGGAGTTTC CTGATTGGG CACTCTCTGT CAAAGCGGTC 660
 ATATGCTGCA GAGTGTCTCC TCTGCAGAAG TCTGAGATAG TGGATGTGGT GAAGAAGCGG 720
 GTGAAGGCCA TCACCTCTGC CATCGGAGAG GCGGCCAAG ATGTGCGGAT GATCCAGACA 780
 GCGCCAGTGG GTGTGGGAAT CAGTGGGAAT GAAGGCATCG AGGCCACCAA CAACTCGGAT 840
 TAGCGCATCG CACAGTTTTT CTACTTAGAG AAGCTTCTGT TGGTTCATGG AGCCTGGAGC 900
 TACAACCGGG TGAACCAAGT CATCTTGATG TGCTTCTATA AGAAGCTGGT CCTGTATATT 960
 ATTGAGCTTT GGTTCGCTTT TGTAAATGGA TTTTCTGGGC AGATTTTATT TGAACGTTGG 1020
 TGCATCGGCC GTTACAATGT GATTTCACCC GCTTTGCCGC CCTTCACTCT GGAATCTTT 1080
 GAGAGGCTCT GCACTCAGGA GAGCATGCTC AGGTTTCCCC AGCTCTACAA AATCACCAGC 1140
 AATGGCGAAG GCTTCAACAC AAAGGTTTTT TGGGTCACG GCATCAACGC CTGTGTCAC 1200
 TCCCTCATCC TCTTCTGGTT TCCCATGAAA GCTCTGGAGC ATGATACTGT GTTTGACAGT 1260
 GGTCTGCTA CCGACTATT TTTTGTGGA AATATTGTT ACACATATGT TGTGTGTTACT 1320
 GTTGTCTGA AAGCTGGTTT GGAGACCACA GCTTGGACTA AATTCAGTCA TCTGGCTGTC 1380
 TGGGGAGACA TGCTGACCTG GCTGCTGTTT TTTGGCATCT ACTCGACCAT CTGCGCCACC 1440
 ATTCCATTGG TCCAGATAT GAGAGGACAG GCAACTATGG TCCTGAGCTC GGCACACTTC 1500
 TGTTGGGAT TATTCTCTGT TCTTACTGCC TGTTTGATTG AAGATGTGGC ATGGAGAGCA 1560
 GCCAAGCACA CTTGCAAAAA GACATTGCTG GAGGAGGTGC AGGAGCTGGA AACCAAGTCT 1620
 CGAGTCTGGG GAAAAGCGGT GCTGCGGGAT AGCAATGGAA AGAGGCTGAA CGAGCGCGAC 1680
 CGCTGATCA AGAGGCTGGG CCGGAAGAGC CCCCAGAGC TGTTCGCGGG CAGCTCCCTG 1740
 CAGCAGGGCG TCCCGCATGG GTATGCTTTT TCTCAAGAAG AACACGGAGC TGTTAGTCAG 1800
 GAAGAATCA TCCGTCCTTA TGACACACCC AAAAAGAAAT CCAGGAAGAA ATAAGACATG 1860
 AATTTTCTG ACTGATCTTA GAAAAGAGAT TCAGTTTGTG GCACCCAGTG TTAACACATC 1920
 TTTGTCAGAG AAGACTGGCG TCCAAGGCCA AAACACCAGG AAACACATT CTGTGGCCTT 1980
 AGTTAAGCAG TTTGTAGTAT ACATATTCCC TCGCAACCTG GAGTGCAGA CCACAGGGGA 2040
 AGCTATCTTT GCCCTCCCAA CTCGTCTGCA GTGCTTAGCC TAACTTTGTG TTAGTCTGTT 2100
 ATGAAGCATT CAACTGTGCT CTGTGAGGTC TCAAATTAAT AACATTATGT TTCACCAATA 2160
 AGAAAAAAA AAAAAA

Seq ID NO: 222 Protein sequence:
 Protein Accession #: NP_057613

1 11 21 31 41 51
 MSVIVTPSG RLRLYCKGAD NVIFERLSKD SKYMEETLCH LEYFATEGRL TLCVAYADLS 60
 ENYEYEWLKV YQEAETILKD RAQRLEECYE IIEKNLLLLG ATAIEDRLQA GVPETIATLL 120
 KAEIKIWLVT GDKQETAINI GYSCRLVSNQ MALILLKEDS LDATRAAITQ HCTDLGNLLG 180
 KENDVALIID GHTLKVALSP EVRRSFLDLA LSCKAVICCR VSPLQKSEIV DVVKKRVKAI 240
 TLAIGDGDAND VGMQIQAHVQ VGISGNEGMO ATNNSDYAIA QFSYLEKLLL VHGAWSYNRV 300
 TKCILYCFYK NVVLYIIELW PAFVWNGFSQ ILFERWCIGL NVVIFTALPP FTLGIFERS 360
 TQESMLRFPQ LYKITQNGEG FNTKVFNGHC INALVHSLIL FWPFMKALEH DTVFDSGHAT 420
 DYLFVGNIVY TYVVVTVCLK AGLETAWTK FSHLAVNGSM LTVLVFFGIY STIWPTPIA 480
 PDMRGQATMV LSSAHFWLGL FLVPTACLIB DVAVRAAKHT CKKTLLEEVQ ELETKSRVLG 540
 KAVLRDSNGK RLNERDRLIK RLGRKTPPTL FRGSSLQQGV PHGYAFSQEB HGAVSQEEVI 600
 RAYDTTKCKS RKK

Seq ID NO: 223 DNA sequence
 Nucleic Acid Accession #: BC017001
 Coding sequence: 1-394

1 11 21 31 41 51

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

AAOCTGGGC AGGGCGGGG CGGGTCGGGG GGGCCCCGAG GGGCCCCGGC CGAGCGGGG 60
CGGCGAGGGC GGCAGCATCC ACTCGGGCGG CATCGCGCGG GTGCACAAGG TGCCGCTGAG 120
CGTGTCTATC CGGCGCTGCG GTTCGGTGTG GGACCCCGCC AAGGTGCAGA GCCTCGTGGA 180
CAGATTCGGG GAGGACCCAG ACAGCGTGCC CCCCATCGAT GTCTCTGGA TCAAAGGGGC 240
CCAGGGAGGT GACTACTTCT ACTCCTTTGG GGGCTGCCAC CGTACGCGG CCTACCGCA 300
ACTGCAAGCA GAGACCATCC CGGCCAAGCT TGTCCAGTCC ACTCTCTCAG ACCTAAGGGT 360
GTACCTGGGA GCATCCACAC CAGACTTGCA GTAGCAGCCT CCTTGGCACC TGCTGCCACC 420
TTCAAGAGCC CAGAAGACAC ACCTGGCCTC CAGCAGGCTG GGCATGCGA AAGGGATAGC 480
AGGGGTGCAT TCTCTTTGCA CCTGGCGAGA GGGTCTGACT CTGGGCACCC CTCTACCGGG 540
CTACAAGGCC TTGGACTCAC TGTACAGTGT GGGAGCCCCA GTTCCACCT CTGTGACAAT 600
AGGATCATGG CCTTACGCTT GAAGCATTAC CGAGAAGGAG AACAGAGATG GGCTTGAAGA 660
GCCACGTGCT GCCGGCTCCA AATTCCCAAG GACAAGGATC CCTCTGCATT TTTGTCTATG 720
TAACCTCTTA TAGGACTAC ATTCACTGCA AAGGAAAGGA AAACCTTGAT TGCAGTGGTT 780
TATTGGTTCA ACAACTCCAC GGAGGTAGGG GTCACTCTTT GGATCTTTT GCCTTAATCT 840
CAGTGTCTGT TACTTCTATG TCCCAAGATG GCTGTCTGAT CCCCAGAAAT CATGTCTGCG 900
TTCAAGGAAG GAGGGGTGGA GGAAGAGGAA GGGCCAAACT AGCTGGACCC GTCACTTCT 1020
ATCAGAAGT AAACCTCGT CAGAAGTCTG TTTCTGTCTC TCTCCCTCTG CATATCTTCA 1080
CTTAGATGCC CTGGGCCCCA GCCAGCTACC ATTGCACCTC TAGCTGCAAA CAAAGCTAAG 1140
ACAGCAGGGA ACAGAAATTG CATGGCTGAA TAGACCAATC GTGTTCATC TACTGAGACT 1200
GGCACACTGC CTCTCTCAAT AAAACTGGGA TCCCATTACC AAGAGAGAAA TGCAGAAATTG 1260
TGTACCAAGT AGCTTTTGTG GTGTAACAAA CCATCCCCAA ACTTGGCAGC TAGAAAACAA 1320
CCCTGTATTT TCCCACAATC CTATGGGTG GCAATTGGG CTGGGCTCAA CAGGGCAGTT 1380
CTGCTGCTCA CACCTGGGAT CCTCATGGA GCTAAGGTCA GCTGTACCT CAGCTGGGCC 1440
TGGATGTTCT AGGATAGCCT TACTCACTTG CCTGGCAGGT GACAGGCTGT TGGCTGGAAT 1500
TGCTTGGTTC TCCTCCATGT GGCCTCTCCA GCAGGCTAGC TCAGGCTTAT TCACATGATG 1560
GCTTCAGGAT TCCAAAGAGA GTGAGAGTAG AAGCTGAAAG ACTTCTTGAG TTCTTGGCCT 1620
GGAACCTGGG CTAGGACAGT GTCACTTCTG CTAAGTTCCT TTGGTCAGAG CAAATCACAA 1680
GGCTTTACCC AGATTCAAGG GATGAGAAAC AGACTACATG TCTTGATGAG GGGAAACCACA 1740
AAGAGCTTGT GGCCTTTTGT CACCTATCAC AAATAATTT GGATGGGTAT TTAITTTGGAT 1800
AAAGGTATTT CCCTCTTCCC CCTTCTCTC TGTCTCATGG GGCCTCACTC TGCCAAGTTG 1860
GAAGGCACTA AGACATTGTC CTGGCCCTCA GGGTCTAGGG GAAGAGGTGT TGGGGCAGGA 1920
AGTGTCTCTC TCCATGGGCT GGACCCACTG TAGTAGGAGT GCCTCCTGT CTGCACTGCT 1980
GGTATGGGCT TAGGCCAGGT AGGCATTCC AGAGGGGCTT CTGAAACCA AGAGTCCCTG 2040
GGGAAAGGGA GAGGAGTAAG GCAGGCTTTC TTCTCACTGC CCTCTAAGGG AACTTGGTCA 2100
CTCGGCATCT TTAAGCCTCA GTTTCTCCAG TTCAATAATA AGGACAAGAG CTTTCCCAT 2160
GCATTCTCTT TCCCGGGGAA AGTTGACTGA GGTGACCACT AATAGAAATT AAAAGGGAGA 2220
GTGTCTTCAG TGCAATGGG CATCTGGAT TGGGTCTTGG AACAAAACA GGACATTAGT 2280
GGGAAAATTG GAAATCTGAA AAAAGTCTGA ATTTAGTTA ATATACCAAT TTCAGTCTCT 2340
TGGTTTTCAG AGATGTACCA TGGTGATGTA AGATGTTGAC CTGGGGTAG GCTGGGTGAA 2400
GGGTATACAG GAATCTTTG TACTATCTCT GCAACTTCTC TGTAAATCTA GTATCATTC 2460
AAAATAAAG TTTATTTAAT TTAATAAAAA AAAAAAAAAA AA

Seq ID NO: 224. Protein sequence:
Protein Accession #: AAH17001.1

50
55
60
65
70
75
80

1 11 21 31 41 51
TLGRAGAGRG APEGPGPSGG AQGGSIHSGR IAAVHNVPLS VLIIRPLPSVL DPAKVQSLVD 60
TIREDDPSVP PIDVLWIKGA QGGDYFYFSG GCHRYAAYQQ LQRETIPAKL VQSTLSDLRV 120
YLGASTFDLQ

Seq ID NO: 225 DNA sequence
Nucleic Acid Accession #: NM_021048
Coding sequence: 1..1110

60
65
70
75
80

1 11 21 31 41 51
ATGCTCTGAG CTCCAAGCG TCAGCGCTGC ATGCTCTGAG AAGATCTTCA ATCCCAAAGT 60
GAGACACAGG GCCTCGAGGG TGACAGGCTC CCGCTGGCTG TGGAGGAGGA TGCTTCATCA 120
TCCACTTCCA CCAGCTCCTC TTTTCCATCC TCTTTCCCTC CCTCCTCCTC TTCTCTCTCC 180
TCCTCTGCTC ATCCTCTAAT ACCAAGCACC CCAGAGGAGG TTTCTGCTGA TGATGAGACA 240
CCAAATCCTC CCCAGAGTGC TCAGATAGCC TGCTCTCTCC CCTCGGCTGT TGCTTCCCTT 300
CCATTAGATC AATCTGATGA GGGCTCCAGC AGCCAAAAGG AGGAGAGTCC AAGCACCTTA 360
CAGGTCTCTC CAGACATGTA GTCTTTACCC AGAAGTGAGA TAGATGAAA GGTGACTGAT 420
TTGGTGCAGT TTCTGCTCTT CAAGTATCAA ATGAAGGAGC CGATCACAAA GGCAGAAATA 480
CTGGAGAGTG TCATAAAAAA TTATGAAGAC CACTTCCCTT TGTGTTTAG TGAAGCCTCC 540
GAGTGCAATG TGCTGGTCTT TGGCATTGAT GTAAAGGAG TGGATCCAC TGGCCACTCC 600
TTTGTCTTGT TCACCTCCCT GGGCTCACC TATGATGGGA TGCTGAGTGA TGTCCAGAGC 660
ATGCCCAAGA CTGGCAATTCT CATACTTATC CTAAGCATAA TCTTCATAGA GGGCTACTGC 720
ACCCCTGAGG AGGTATCTCT GGAAGCACTG AATATGATGG GGCTGTATGA TGGGATGGAG 780
CACCTCATTT ATGGGGAGCC CAGGAAGCTG CTCACCCAAG ATTGGGTGCA GGAAAACATC 840
CTGGAGTACC GGCAGGTGCC TGGCAGTGAT CTGCACCGT ATGAGTTTCT GTGGGGTCCA 900
AGGGCTCATG CTGAAATTAG GAAGATGAGT CTCTGAAAT TTTTGGCCAA GGTAAATGGG 960
AGTGATCCAA GATCCTTCCC ACTGTGGTAT GAGGAGGCTT TGAAGATGA GGAAGAGAGA 1020
GCCCAGGACA GAATTGCCAC CACAGATGAT ACTACTGCCA TGGCCAGTGC AAGTTCTAGC 1080
GCTACAGGTA GCTTCTCTTA CCTGAATAA

Seq ID NO: 226 Protein sequence:
Protein Accession #: NP_066386

85

1 11 21 31 41 51
MPRAPKQRRC MPEEDLQSQS ETQGLEGAQA PLAVEEDASS STSTSSSFPS SFPSSSSSSS 60

WO 02/086443

PCT/US02/12476

SSCYPLIPST PEEVSADDET PNPPQSAQIA CSSPSVVASL PLDQSDGESS SQKEESPSTL 120
QVLDPSSSLP RSEIDEKVTD LVQFLFKYQ MKEPITKAEI LESVIKNYED HFPLLFSEAS 180
ECMLLVFGID VKEVDPTGHS FVLVTSGLLT YDGMLSDVQS MPKTGILILI LSIIFIEGYC 240
TPEEVIWEAL NMGLYDGMG ELIYGEPRKL LTQDNVQENY LEYRQVPGSD PARYEFLWGP 300
RAHAEIRKMS LKFLAKVNG SDPRSFLPLY EEALKDEEER AQDRIATDD TTAMASASS 360
ATGSFSYPE

Seq ID NO: 227 DNA sequence
Nucleic Acid Accession #: NM_005025.1
Coding sequence: 82-1314

1 11 21 31 41 51
GCGGAGCACA GTCCGCCGAG CACAAGCTCC AGCATCCCGT CAGGGGTTGC AGGTGTGTGG 60
GAGGCTTGAA ACTGTTACAA TATGGCTTTC CTGGACTCT TCTCTTGTCT GGTTCGCAA 120
AGTATGGCTA CAGGGGCCAC TTCCCTGAG: GAAGCCATTG CTGACTTGTC AGTGAATATG 180
TATAATCGTC TTAGAGCCAC TGGTGAAGAT: GAAAATATTC TCTCTCTCC ATTGAGTATT 240
GCTCTTGCAA TGGGAATGAT GGAAGTTGGG GCCCAAGGAT CTACCCAGAA AGAAATCCGC 300
CACTCAATGG GATATGACAG CCTAAAAAAT: GGTGAAGAAT TTTCTTTCTT GAAGGAGTTT 360
TCAACATCGT TAACCTGCTAA AGAGAGCCAA TATGTGATGA AAATTGCCAA TTCTTGTITT 420
GTGCAAAATG GATTTCATGT CAATGAGGAG TTTTGTCAA TGATGAAAA ATATTTTAAT 480
GCAGCAGTAA ATCATGTGGA CTTCAGTCAA AATGTAGCCG TGGCCAACTA CATCAATAAG 540
TGGGTGGAGA ATAACACAAA CAATCTGGTG AAAGATTGGG TATCCCCAAG GGAATTTGAT 600
GCTGCCACTT ATCTGGCCCT CATTAAATGCT GTCTATTTC AAGGGAACTG GAAGTCGAG 660
TTAGGCGCTG AAAATGCTAA AACCTTTTCT TTTCACTAAG ATGATGAAAG TGAAGTCCAA 720
ATTCCAATGA TGATACAGCA AGGAGAATTT TATTATGGGG AATTTAGTGA TGGCTCCAAT 780
GAAGCTGGTG GTATCTACCA AGTCTAGAA ATACCATATG AAGGAGATGA AATAAGCATG 840
ATGCTGGTGC TGTCCAGACA GGAAGTTCCT CTGCTACTC TGGAGCCATT AGTCAAAGCA 900
CAGCTGGTGG AAGAAATGGG AAACCTCTGT AAGAAGCAAA AAGTAGAAGT ATACCTGCC 960
AGGTTACAGT TGAACAGGAA AATTGATTTA AAAGATGTTT TGAAGGCTCT TGAATAACT 1020
GAAATTTTCA TCAAGATGC AAATTGACA GGCCTCTCTG ATAATAAGGA GATTTTCTT 1080
TCCAAAGCAA TTCACAATC CTTCCTAGAG GTTAATGAAG AAGGCTCAGA AGCTGCTGCT 1140
GTCTCAGGAA TGATTGCAAT TAGTAGGATG GCTGTGCTGT ATCTCAAGT TATTGTGAC 1200
CATCCATTTT TCTTTCTTAT CAGAAACAGG AGAAGCTGGT CAATCTTATT CATGGGACGA 1260
GTGATGATC CTGAACAAT GAACACAAAT GGACATGATT TCGAAGAACT TTAAGTTACT 1320
TTATTTGAAT AACAGAGAAA ACAGTAACTA AGCAGATATT GTTTGCAACT GGTATATATT 1380
TAGGATTTGT GTTTTACAGT ATATCTTAAG ATAATATTTA AAATAGTTCC AGATAAAAC 1440
AATATATGTA AATTATAAGT AACTGTCAA GGAATGTTAT CAGTATTAAG CTAATGTTCC 1500
TGTTATGCA TTGTGTTGT GTGCTGTGT TTAATAAAG AGTACCTATT GAACATGT

Seq ID NO: 228 Protein sequence:
Protein Accession #: NP_005016.1

1 11 21 31 41 51
NAFLGLFSLV LQSMATGAT FPEEAIADLS VNMYNRLRAT GEDENILFSP LSIALAMGMM 60
ELGAQGSTQK EIRHSMGYDS LKNGEESFSL KEFSNMVTAK ESQVVMKIAN SLFVONGFHV 120
NEEFLQMKKK YFNAAVNHVD FSNVAVANY INKWENNTN NLVKDLVSPR DFDATYLYAL 180
INAVYFKGNW KSQFRPENTR TFSPTKDES EVQIPMYQQ GEPYGEFSD GSNEAGGIYQ 240
VLEIPYEGDE ISMMLVLSRQ EVPLATLEPL VKAQLVEEWA NSVKKQKVEV YLPRFTVEQE 300
IDLKDLVKAL GITEIFKDA NLGLSDNKE IFLSKAIHKS FLEVNEEGSE AAASGMIAI 360
SRMAVLYPQV IVDHPFFLI RNRRGTGILP MGRVMPHPTM NTSGEDFEEL

Seq ID NO: 229 DNA sequence
Nucleic Acid Accession #: NM_003695
Coding sequence: 12-398

1 11 21 31 41 51
CGACATCAGA GATGAGGACA GCATTGCTGC TCCTTGACG CCTGGCTGTG GCTACAGGGC 60
CAGCCCTTAC CCGCGCTGC CACGTGTGCA CCAGCTCCAG CAACTGCAAG CATCTGTGG 120
TCTGCCCGGC CAGCTCTCGC TTCTGCAAGA CCAGCAACAC AGTGGAGCCT CTGAGGGGGA 180
ATCTGGTGAA GAAGGACTGT GCGGAGTCGT GCACACCCAG CTACACCCCTG CAAGGCCAGG 240
TCAGCAGCGG CACCAGCTCC ACCCAGTGCT GCCAGGAGGA CCTGTGCAAT GAGAAGCTGC 300
ACAAACGCTGC ACCCACCGCG ACCGCCCTCG CCCACAGTGC CCTCAGCCTG GGGCTGGCCC 360
TGAGCCTCCT GGCGTGCATC TTAGCCCCCA GCCTGTGACC TTCCCCCAG GGAAGGCCCC 420
TCATGCCCTT CTCTCCCTTT CTCTGGGGAT TCCACACCTC TCTTCCCCAG CCGGCAACGG 480
GGGTGCCAGG AGCCCCAGGC TGAGGGCTTC CCCGAAAGTC TGGGACCAGG TCCAGGTGGG 540
CATGGAATGC TGATGACTTG GAGCAGGCCC CACAGACCCC ACAGAGGATG AAGCCACCCC 600
ACAGAGGATG CAGCCCCCAG CTGCATGGAA GGTGGAGGAC AGAAGCCCTG TGGATCCCCG 660
GATTTCACAC TCCTTCTGTT TTGTGCGGT TTATTTTGTA CTCAAATCTC TACATGGAGA 720
TAAATGATT AAAC

Seq ID NO: 230 Protein sequence:
Protein Accession #: NP_003686

1 11 21 31 41 51
MRTALLLLAA LAVATGPAIT LRCHVCTSSS NCKHVVCPA SSRFCKTTNT VEPLRGNLVK 60
KDCAESCTPS YTLQGQVSSG TSSTQCCQED LCNELHNAA PTRTALAHS LSLGLALSLL 120
AVILAPSL

Seq ID NO: 231 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 126-752

1 11 21 31 41 51
 5 CCGGGCAGGT GGCCTCATGT CGGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCTTGGAGC 60
 AGGGCGGCAG GAATTCTGAT GTGAACTAA CAGTCTGTGA GCCCTGGAAC CTCCACTCAG 120
 AGAAGATGAA GGATATCGAC ATAGGAAAG AGTATATCAT CCCCAGTCTCT GGGTATAGAA 180
 GTGTGAGGGA GAGAACCAGC ACTTCTGGGA CGCACAGAGA CCGTGAAGAT TCCAAGTTCA 240
 GGAGAAGTCG ACCGTTGGAA TGCCAAGATG CCTTGGAAAC AGCAGCCCCG GCCAGGGGCC 300
 TCTCTCTTGA TGCTCTCATG CATCTCTCAGC TCAGAAATCCT GGATGAGGAG CATCCCAAGG 360
 10 GAAAGTACCA TCATGGCTTG AGTGCTCTGA AGCCCATCCG GACTACTTCC AAACACCAGC 420
 ACCCAGTGGG CAATGCTGGG CTTTTTCTCT GTATGACTTT TCGTGGCTT TCTTCTCTGG 480
 CCGGTGTGGC CCACAAGAAG GGGGAGCTCT CAATGGAAGA CGTGTGGTCT CTGTCCAAGC 540
 ACGAGTCTTC TGACGTGAAC TGCAGAAGAC TAGAGAGACT GTGGCAAGAA GAGCTGAATG 600
 AAGTTGGGCC AGACGCTGCT TCCTTGCGAA GGGTTGTGTG GATCTTCTGC CGCACCAGGC 660
 15 CATCTGTGTC CTGATGATCA CGCAGCTGGC TGGCTTCAGT GGACCAAAAT 720
 TTCAGGATGG CTGTATTCTG CGGTGAGAAT GAGAGAGTCA AGCTGGGCGG AATCTCTCGC 780
 CAAGAGTCTA GCCTTCTCTT GGAGACTGCT CCATCAGTGC CGAGGTGTGT GGGAACAGGC 840
 TTCAGTGCAC CGCCATCTTA CTGAGTTGCT TCACGTGAGG AAAAGGGGGC TTTGGCCCTG 900
 TGACTCAGTT CCACATTTTG GATTGCATAC TGGAAAAGAA GCCAATCTTC TTGCTAGTAA 960
 20 ACCAGCAACC CCGCTGTATA CAGTGGTGAC CCAAGCAATG GATATAAACC TAAAAATCTG 1020
 AGGGAGGGGA GAGGTGGAAAT ACAGTAGTTC TTGGAATCTG AAGTCTCCTA TTTGATCAGG 1080
 TTATTTCCTG GACTTTGGCA AAAATCTGAT TGGTGGGGAT CTCTTAGGAC CTAGTGAGCA 1140
 TCTGGTATTA ATTTAATCTC AGGAAAAACA AGAAATTAAC CCAGAGAGAG TCTGGGTTTT 1200
 GGAATTGAGC GTAGCTAGCT CCAGACCGTG GTGTCTGGCC TCCATTTTGT TCTGTCTTCT 1260
 25 AGCTCTGACT TACAGCTGCA GTCACTTTTG CTATAAGGCA CCTGGGTAGA AGGGTGGATG 1320
 GGCTTCACAT CAATTTTTTT CTCTCTTAG GGTGGGGGAT TGGTTTGCT TCTTTTGT 1380
 GTGGTTTTTT GTTTTTTCTT TGTCAAGATT GATTTTTAGA TGCAAGGACT TGAAGAAGCC 1440
 CAGAAGGATG CCACCAGTTT TTCCTTGAGG CCTAGGATTT TTTATTCTGT CCGAGCAGA 1500
 GGTAAATCCT CACAACCTTG TGCACCAGTA GCACCAGCCA TTTTGGCAG AGTACTCTTT 1560
 30 TGGGAGCTT TCGCTTTTGT TTTGTTTTTA ATTCTCTTTC CTTAGCAGCA AGGTCTTTTT 1620
 TCCTAGAGAA TCTACTCCGT TGCAGAACTA TTGCAACCTC AGGAGCCCTC ACTGATTGAG 1680
 TGCTGTGAGC CTGATATACT ACTTTGGACT CTGAAAACAG ATATGGGTTT TATTCTCTAT 1740
 TTCTACTGTG TGTGTTTAAA CAACCGTCGG AGACCAGATG ACCTGTTAGA TGGCTAGTCC 1800
 TGTATACTC GACTCTGTAT GTTCAATGT ATGTACTGCA AATGCTTCAC CTGCTGTACA 1860
 GTGTTTGTGA GATGCTCTTT GAAGATGGTA CTTTTATATT T

Seq ID NO: 232 Protein sequence:
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 40 MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRP TRPLECODAL ETAARAEGLS 60
 LDASMHSQRL ILDEHPKPKG YHHGLSALKP IRTTSKHQHP VDNAGLFSCM TFSWLSLAR 120
 VAHKRGELSM EDVWSLSKHE SSDVNCRRLE RLWQEELENEV GPDAASLRRV VNIQFCTRLI 180
 LSIIVCLMITQ LAGFSGNPFQ DGCILRSE

Seq ID NO: 233 DNA sequence
 Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51
 50 TTTTAATGGT GCTCATATAT ACTGTATTTT TGTGTGTTTA GTTTTACTTA TTGAGAGTGT 60
 CACAACATGA ATCACAATAA CATGATTTTT TTTTTTACT TTTACTCCCC AAATTATTCA 120
 TGTTCCTTAG ATCGTAGTCA TTGAGAAGTC CCAATAACTC TAAACTTTTG AGTTATAACG 180
 55 TAGTAAACTT CTCTTTCATC TTTGTGTTAG CTCTGTAGTC TTAACCTGGA TTTTAATTTT 240
 TTTGTTTCCA AAGTCACAAT TGAATTATTC TTAGATACCT TAAGCCACTG AATTCTAGTT 300
 TGTTTGATCG AAGCAAAAC AAGGTGACAG TTTATTTTCA AACACTAATC TCTGTATATT 360
 TTGTTATGGT ATATCTTTTT ATTAATAATT TATTTTGACT AAGCTTTCAT AAAATATTG 420
 AAGCTATTTT AATCATCAAG TATGGAAAAC AAATTACTAT TGCAATTTCC TATATATGCA 480
 60 TATATTATGG ATTAACGAGA ATTGTATCAT TTTTGGCCTA ATGTCTGGAT ATAAAAGATA 540
 ATTAGCTTAC TATAGTATTA ATAAATTTT CAGTTGGTTT GGGCAAAATT AAACCTGAAA 600
 AATAGGTTAA AAGTAGTTAA CAAATTAAAC TTAATAATT ATACCTGATT TTTTTCTG 660
 AATTAAGTA CATTTTAAAT GAGCTTTATA ATACCTTAAA AAGTTGGTTC TAAATTAATA 720
 TATGAAAGCT CTGGCTATCA TCCGGGATA GTAATTTCTA ATTATATAGT ATTTCAAAC 780
 65 TATATATTTT TTAGTTCCCT TGAGATAACT AATTTCTAAT TATATATGTT TCAAAAACCA 840
 TATCCTGTAT TTTTTTAAG AATTGTTTAA TAAATAGTGC ATAAGATACA AGGTCTGCAT 900
 TAGAAGACCC ACTCTTACTA GGTCCCTAA GATCTGCCA TAGATTTTTT TTTTTTTTT 960
 TTTTTTTTAG GTAGTTTAAA GCAAGCACTG ATACCACTGG GAGTTGGTCT TGATCTAGGA 1020
 GATTCTGTTA AGCATCCAAA AACAAATGCT AATTTCAAGT CTAGGTTAT GGCTTGTGAC 1080
 70 TCCAGATAAA AGATGGAGAA TACCTCATGT ACTGTGACTT GAAATGAAT TCTTAAAAAT 1140
 CTTAGGCTCT CTCCATGTAT CTTCTTAA GAAAAGTTTC TGAGTGTGAT CTCTCTTTG 1200
 CCATAGTATC AAGTGGAGGG TAGTTCAGAA AAGTTAATAG GAAATCTTTT GTGACAGCAG 1260
 ACTATAATAG AAGTTTGAGT AATATTTTAA TAAATTTATA TAATCAAAT GATAAAAATG 1320
 TATCAATGTT ATCCAATGAT TTTTATTA AAATTACCTT ATTATTAGAA CTGTGCCTAT 1380
 75 TACATAAAAA GTGCTCATGT ATTGAATTT TAAATAATTT ATTTAAATCA AGACCACCAT 1440
 AAGTCATTAA TAAATTAATA ATTGTTTAA ATCAGTGGTT TTCAACCTC ACTTCATATT 1500
 AGAATCATCT GAGGACTTTT AATATGGAAT CCACCTCATA ACAATTAAGT CTAAATTTCT 1560
 GGAAGATGGA GCCATGCTTG TTTTCCAAA AGCTCTTTGA GTGATTCTAA TTTGTAGTCA 1620
 GAGTTGAAGA CCATGCTCTT AAATTAGTGC AGGAAAATGC TTTTATTCT CCATGTTTAA 1680
 80 CTTTAAAC TAGTATGTA CCAAGTTAAG TTTGATGGT TTAATTTCCA CTAAAGAAC 1740
 TATTCTCTTA ATAAGTACGA TTTATTACAT GAAATTTAAG AGTTTAAAGT CCATCAAAT 1800
 AGCCCTTGTG TAAGATTATT ATTTCTCTC TATAACTTCA AAATAGATAT TTCATTCAA 1860
 CTGTTCAAGT GAGAAAACAT AATGATTTT TTTTCTTTC CTCTGGAGCT GCCGTGTCAG 1920
 TGAGATGGAG GAGGTGGGCA CATTTAAGGT CAGTTCACCT ACCTATGGTT CAGAGTTCTG 1980
 85 ATCATATGGA AGTTTGGAAA AGAGAGCTTA TCACAGGTTT GTATGCTGGT GAATGGATAG 2040
 TTTTAATTTCT CACTGTCTCA AAAGAGAATC AGCTCTCCAG CAGTTCTAGA AAAGCTTTGA 2100
 CAATCCCAA GGGGAGTGT TACCTTACTC CTTCAGTCT TCTTAGAAGG TAGAATTAAG 2160
 TTTCTGGAAT TGCACTTACA TGTTTCTTA TTAACATTCA GAATGGGAA TATTAATTTT 2220

TCCAGTGAGT AGTTTCTGA AATTGGTAAC TTGGAGAGTA AAATAACGTA TTTTGCTTTT 2280
 CAATTTTGTG TTTGTTTACT TTTATGTAAA AATTGTATAT GTGAATTACA CAGTTCTAAT 2340
 AAAACCTCAT GCCTTTTCAT TACATCTAAT TTGAACCTCT AACTTCAGTG CCAGAAAGTGC 2400
 TTTAAGATG CTTTAATGAA AAGTATTAAAG AAAATATATA GATTGTATG TCAGTTTATA 2460
 CTTGAGAAAT CCAATATGAT GTCATATTTA TTTTGTAGA AACCTCCTAA TTGGATAACT 2520
 AGATGGTATT TAAATGAAT GCCCAAAAT ATCTGTGACC TTTGTCCAAA AGTTTATCTG 2580
 TTGGAAGCCG CCAGCCATTC ATGTAGAGAG TTTATAAGAA AATAATTTAA AATTGTATGC 2640
 ATTTTATATT ACTATGGTAT CTGTGTACCA TATTCTAAG TATTCAATTAT TAAATTGGTA 2700
 CTTCTTAAAA CCATAACCTG GCTTGCCCTT TAGTGTATA CACAAAATCC AACATTGTAT 2760
 ATAGAGATTG TTCTTTTATG AAGAAGAGCT GACGTAATTT ATTACCAAGT CATCTGCACA 2820
 AAGACATTAA CATAAGTCTC TGAGCAGTGA TACATTTTCA AACATGAAGA GTGACAAACA 2880
 CCACATTAAA CAACCCAGCG AACACTCAGA CTTGGCACTT TCCTACGAAT CCATCCTATA 2940
 TGTGCTGGT ATGCGCTCTG GCATAACTTA CACGAATCGT CCTCCCTACT TGTCTACGCT 3000
 CCTTCATCAA GCATCTGCCA ACACATTAC CTCTAACTTG TACAACCTTA CCAACTCACC 3060
 ACAACATCTG CCACTCTATC CTATCAACTG CCAACCTTAA GACCCCAAC ACAACACAC 3120
 CCCCACACAC AAAACCACTA AATCATAACC ACCACACACG CCACACACCA CACACCCACC 3180
 CACACACACA ACACACACG ACCAACACCC CCACCAAAA CAAGCTAACA ACCACAAACA 3240
 GACACACAT CACATACACT CACTACCCCC CCATACTCCC ACCCACCA

Seq ID NO: 234 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 27-281

1 11 21 31 41 51
 AGCAGGAGGA GAGCTGGCGG GAAGACATGC ACCCCTTGAA GACCCAGAGA GAGGCGCTCT 60
 GTCTACCGCG TAGCAGTTAC ATCAGACTGA GACACTTCCT GTTTACAGGA GACTATAAAA 120
 TTCTGCCCCC GTGCTCATTT GGGGCTGACG CCATTTTAGG CCTCAGCCCA TCTGCACCCA 180
 GGGCGTCACT GAAACAGTGT GTTGCTCCAC ACCGCTTGTG TTTGCTTGTG GGGCGGCTCT 240
 CAGGGTTCCG ACCAATCCAA GAGCCTTGCA GAAAGCATTG ACGTGCTTTT CTCTTTGGCA 300
 GAGTTTTTCT TTGCTCTGAT CTGSGAGACA TCCCTCTGCC TAGTGGAAAC ATAAGGAATA 360
 CAGAAAGAAAT GCAAGGAGAT AGACCAACGT GAGATTCTCC TTCATGCACT CAAGAGAAAG 420
 ATGTTGCAGG AAGAGCTAGT CTTTCAGGCT GGGCTGGTGA CCTGAGAAAG AATGTCCAGC 480
 TTTTCTTCTC CACTTGGCAT ATCAAGAGCC AGGCGTGGAA GACTAAAACA GGAATGTGTT 540
 ATAAAAACCTG TTCAGCGGTT CGCCAACAAG AAGTGGTAAA GTAGCAAAAA TGGGGATGGA 600
 GATGCCAGGA GGAAGATGTC CAGGGGTAAA GTGGGAAAT GGGAACTGA AGCCAGGAGG 660
 TCAAGCCAAAG CCAACAGGTG TTCTGTTTTT CATCACAGAA CTAAATAGTG GTGCTGAGGA 720
 CTCAAACCCG GGGAGGCCCA CTCTAGAACC CATGCTGGTC ATCCATATCC CCAAGGCCCT 780
 GGTGAGAAAC CAGCTAAGCA GATGGCTTGG GTCATCAGGA CGTCCATTAC ATCCAAAGGA 840
 AGACAGCCTG TGACGTTTCA AAAGCAAAAG TCCCTACCA GCCAGTGAAG CTACCTGATT 900
 TCTCAGTATC TTACGCCAG TGACACGATC TACCCTCAA ACTTAAAAAA AAAAGGGAAA 960
 CATAAACACA TAACAGCAGC AGCAATAAIT AAAGATGAGA TGAGAACAA TAAGAAAAAA 1020
 GGAAGGCTCT CCGTGACTG TTTTATTTT AGGGAAACAG AGAGGAAGAA GAATGATTTT 1080
 TCTTTGATG ACTCTATATC CAACTCTGAG GTTTGATTAA AGAAATGACC TTGAACCACA 1140
 GCAAGAGAAA ATAAAGACA ATTTCAGTA AGTATGCCAG TTGCAATTAA TGATTTACTT 1200
 TTTATTTTAA AACTGAATTC AGCAGAGATT TACATGCATT ACGATGATTA ACATCTGAAA 1260
 TTTGACCTTG AAATAATCTT TACATTGTAA ATTCTTAATG ATCAAAACAA GGTCTCAGT 1320
 GATTAAACACA TATTAGTAAT TAATTATTAA AGGAGAATAA TTGCAATAC AACATTCCCTA 1380
 AAATCTCAAG GCTTTTAAAG CATTTGTACA AATGACTGGA CATTTTAAAT ATTTGAAAAA 1440
 AAAAAAAGC CCTCCATCTG ATTCTCATTT TCATTGTCTG TGCAACAAAC AAAAAGGTAT 1500
 GCATCTCTCT TCTCATTTTC CACTGTCTCG CAAGCTAGAA ATTCTCAACA CTACCTTTGA 1560
 TCCCATCAAA GCCAAAGAAA GAAAGAAAA TTGTTCTGTA CAGATATATG ACATTAAAAA 1620
 ATAATCCC

Seq ID NO: 235 Protein sequence:
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 MHPLKTRQEA VCLPRSSYIR LRHFLFTGDY KIPAPCSFGA DAILGLSPSA PRRSLKQCV 60
 PHLVLVGA LSGFRPIQEP CRKH

Seq ID NO: 236 DNA sequence
 Nucleic Acid Accession #: NM_002075
 Coding sequence: 406..1428

1 11 21 31 41 51
 CCACATAGG GGCAGACCTG TCCATCCTTC TCTGTGGGTC CCTGTACCT TTCTCCCCCA 60
 ACAGGATCAG ACCCAGAGGC AGCTGGTGTG GGTGTGTGCA GAAGAAGGAT TATCCAGATC 120
 AGTCCCTTTC AATCTCAGCT CCTGCTGTGA CCTCCCATTA CTACCAAAC CCTCTCCCCC 180
 ACCACCTTGA GCTGAGGAGC ACAGTTTGAG GCCCCCCCAA CCCCCCGCCG GTCGGGGCCA 240
 GGGCAGGCCA GGCACGCTCC TCTGGCAGCA GAGCCTGGGC AGGTGACGGG CGGGCGGGG 300
 CGTGCAGCT GAGGGAGTAA GGAGGCTCCC AGGAACCGGA GCTGGAACC CGGCCAGGT 360
 CCAGCCAGAG CCCAAGAGCC AGAGTGACCC CTCGACCTGT CAGCCATGGG GGAGATGGAG 420
 CAACTGCGTC AGGAAGCGGA GCAGCTCAAG AAGCAGATTG CAGATGCCAG GAAAGCCTGT 480
 GCTGACGTTA CTCTGGCAGA GCTGGTGTCT GGCCTAGAGG TGGTGGGACG AGTCCAGATG 540
 CGGACGCGGC GGAAGTTAAG GGGACACCTG GCCAAGATT ACGCCATGCA CTGGGCACT 600
 GATTCTAAGC TGCTGTAAAG TGCTCGCAA GATGGGAAGC TGATCGTGTG GGACAGTAC 660
 ACCACCAACA AGGTGCACGC CATCCCACTG CGCTCCTCT GGGTCATGAC CTGTGCTTAT 720
 GCCCATCAG GGAATTTTGT GGCATGTGGG GGGCTGGACA ACATGTGTT CATCTACAAC 780
 CTCAAATCCC GTGAGGGCAA TGTCAGGTC AGCCGGGAGC TTTCTGCTCA CACAGTTTAT 840
 CTCTCTGTCT GCGCTTCTCT GGATGACAAC AATATTGTGA CCAGCTCGGG GGACACACG 900
 TGTGCTTGT GGGACATTGA GACTGGGAG CAGAAGACTG TATTGTGGG ACACAGGGT 960
 GACTGCATGA GCCTGGCTGT GTCTCTGTAC TTCAATCTCT TCATTTCGGG GGCTGTGTAT 1020
 GCCAGTGCCA AGCTCTGGGA TGTGCGAGAG GGGACCTGCC GTCAGACTTT CACTGGCCAC 1080

5
10
15

GAGTCGGACA TCAACGCCAT CTGTTTCTTC CCCAATGGAG AGGCCATCTG CACGGGCTCG 1140
 GATGAGCGTT CCGTCCGCTT GTTTGACCTG CGGGCAGACC AGGAGCTGAT CTGCTTCTCC 1200
 CACGAGAGCA TCATCTGCGG CATCACGTCC GTGGCCTTCT CCCTCAGTGG CCGCCTACTA 1260
 TTGCTGGCT ACAGACGACT CAACCTGCAAT GTCTGGGACT CCATGAAGTC TGAGCGTGTG 1320
 GGCATCCTCT CTGGCCACGA TAACAGGGTG AGCTGCCTGG GAGTCACAGC TGACGGGATG 1380
 GCTGTGGCCA CAGTTCCTG: GGACAGCTTC CTCAAAATCT GGAACTGAGG AGGCTGGAGA 1440
 AAGGGAAGTG GAAGGCAGTG AACACACTCA GCAGCCCCCT GCCGACCCC ATCTCAITCA 1500
 GGTGTTCTCT TCTATATTCC GGGTGCCATT CCCACTAAGC TTTCTCCTTT GAGGGCAGTG 1560
 GGGAGCATGG GACTGTGCTT TTGGGAGGCA GCATCAGGGA CACAGGGGCA AAGAAGCTGCC 1620
 CCATCTCCTC CCATGGGCTT CCCTCCCCAC AGTCTCCACA GCCTCTCCCT TAATGAGCAA 1680
 GGACAACTG CCCCTCCCCA GCCCTTTGCA GGCCAGCAG ACTTGAGTCT GAGGCCCCAG 1740
 GCCTAGGAT TCCTCCCCCA GAGCCACTAC CTTTGTCCAG GCCTGGGTGG TATAGGCGT 1800
 TTGGCCCTGT GACTATGGCT CTGGCACAC TAGGGTCTTG GCCCTCTTCT TATTATGCT 1860
 TTCTCCTTTT TCTACCTTTT TTTCTCTCCT AAGACACCTG CAATAAAGTG TAGCACCCCTG 1920
 GT

Seq ID NO: 237 Protein sequence:
 Protein Accession #: NP_002066

20
25
30

1 11 21 31 41 51
 MGEMEQLRQE AEQLKKQIAD ARKACADVTL AELVSGLEVV GRVQMRTRRT LRGLHLAKIYA 60
 MHWATDSKLL VSASQDGKLI VNSDYTTNKV HAIPLRSSNV MTCAYAPSGN FVACCGLDNM 120
 CSIYNLKSRE GNVKVSRELS AHTGYLSCCR FLDDNNIVTS SGDTTCALWD IETGQQXTVP 180
 VGHGTDCMSL AVSPDFNLFI SGACDASAKL WDVREGTCRQ TFTGHESDIN AICFPNGEA 240
 ICTGSDDASC RLPDLRADQE LICPSHESII CGITSVAFSL SGRLLFAGYD DFNCNVWDSM 300
 KSERVGLISG HDNRVSLGV TADGMVAATG SWDSFLKIWN

Seq ID NO: 238 DNA sequence
 Nucleic Acid Accession #: CAT cluster

35
40
45
50

1 11 21 31 41 51
 TCCCAATGTG TNGAACCTAC CATAAATCTT TTCTTACNG GACAATCTTA TNCTAANCAA 60
 TACCATTTCG TTTTAAGGCA GATAATCCTC CAAGTTTCTT AATGATATCT GAAACTATTA 120
 ACTGATTCTG TGAATTATGA AATCTGAAAA GGAATTGGAA GTTGCTAAAA ATCTATCATT 180
 TGCAATGACC AGTGTGAAGC ACAGTGAAGT GAGAATCGCT GCCCTGACAC CAAAGAAAAA 240
 TAAGTGAAGT GAAAGCTGAA GAATCACCGG CTTCACTGAC ATGGAACCCA GTGATTGTAT 300
 TTTTGACGAG TATCGGGTGA CTTTGAGGTG GTCAAGAAAC CACACTTTAA GAACAATGTC 360
 CAAAAAGGGG AAAAAAAGA GCAACCAAGG AAAAAAATC CATAAATTC CACAGAAGAA 420
 AAGAAAGAAA AATAAAATAC ACATATGGGA CGATGGAGAA AACACAGTTAC ATTTCTTTAT 480
 GGATCAAGAA GTTTGTGTAC ACATAATCTC ATTTTGAGAT ATATACTAT TTTTGTCTTT 540
 CAGAAAGTGA TCAAAATATT TCAAAATGCT GTCTTATGAA ACTACAATAT TCTCACAGAT 600
 TAGAAAAATT TTTCTGTAAG AGTCAGATAG TAAATATTTT AGGTTTTCGA GTGTCTTTTG 660
 CACTACTCTA ACTTTCTCTA TGTAGCACA GAGTAGCTGT GGTACTGTGC AAATAAATTG 720
 CTTGTGTTC AATAAAGCTT CATTTACAAA AACATGCCAT GGGCCATATT TGGCCTGTAC 780
 ACTGTGTTT GCCAAGTCTT AATATAGTTG CTTAGCAAGT ATTGTGAGCT ATTTGAGGAA 840
 GACATGAAAG TCTATTGGGT TGCTAAAAAG TATGTAGAAA TTCAAAGGAA AATTAATAAT 900
 TAGGCTAAGT TATAATACAC TGTTTTAAAC ATTGTAAAAA GTAAGAGAAA TTTACAAATA 960
 AAAATCCCAA ATAAAA

Seq ID NO: 239 DNA sequence
 Nucleic Acid Accession #: NM_001786.1
 Coding sequence: 130-1023

55
60
65
70
75
80

1 11 21 31 41 51
 GGGGGGGGGG GGCACCTGGC TTCAAAGCTG GCTCTTGGAA ATTGAGCGGA GAGCGACGG 60
 GTTGTGTAG CTGCGCGTGC GGCCGCGCGG GAATAATAAG COGGGATCTA CCATACCCAT 120
 TGACTAATCA TGGAAAGATTA TACCAAAATA GAGAAAAATG GAGAAGGTAC CTATGGAGTT 180
 GTGTATAAGG GTAGACACAA AACTACAGGT CAAGTGGTAG CCATGAAAAA AATCAGACTA 240
 GAAAGTGAAG AGGAAGGGGT TCCTAGTACT GCAATTCGGG AAATTCTCTT ATTAAGGAA 300
 CTTGCTCATC CAAATATAGT CAGTCTTCAG GATGTGCTTA TGCAGGATTC CAGGTTATAT 360
 CTCATCTTTG AGTTTCTTTC CATGGATCTG AAGAAATACT TGGATTCTAT CCCTCCTGGT 420
 CAGTACATGG ATTCTTCACT TGTTAAGAGT TATTTATACC AAATCTTACA GGGGATTGTG 480
 TTTTGTCACT CTAGAAGAGT TCTTACAGA GACTTAAAC CTCAAATCT CTGTATTGAT 540
 GACAAAGGAA CAATTAAACT GGCTGATTTT GGCCTTGCCA GAGCTTTTGG AATACCTATC 600
 AGAGTATATA CACATGAGGT AGTAACACTC TGGTACAGAT CTCCAGAAGT ATTGCTGGGG 660
 TCAGTCTGTT ACTCAACTCC AGTTGACATT TGGAGTATAG GCACCATATT TGCTGAACTA 720
 GCAACTAAGA AACCACTTTT CCATGGGGAT TCAGAAATGG ATCAACTCTT CAGGATTTC 780
 AGAGCTTTGG GCACCTCCCA TAATGAAGTG TGGCCAGAAG TGAATCTTT ACAGGACTAT 840
 AAGAATACAT TTCCCAATG GAAACCAGGA AGCCTAGCAT CCCATGTCAA AAACCTGGAT 900
 GAAAATGGCT TGGATTGTCT CTCGAAAATG TTAATCTATG ATCCAGCCAA ACGAATTCT 960
 GGCAAAATGG CACTGAATCA TCCATATTTT AATGATTGG ACAATCAGAT TAAGAAGATG 1020
 TAGCTTTCTG ACAAAAAGTT TCCATATGTT ATGTCAACAG ATAGTTGTGT TTTTATTGTT 1080
 AACTCTTGTG TATTTTGTG TTATATATAT TTCTTTGTTA TCAAACTTCA GCTGTACTTC 1140
 GTCTTCTAAT TTCAAAAAA TAACTTAAAA ATGTAAATAT TCTATATGAA TTTAAATATA 1200
 ATTTCTGTAA TGTGAAAAA AAAAAAATA AAAAA

Seq ID NO: 240 Protein sequence:
 Protein Accession #: NP_001777.1

85

1 11 21 31 41 51
 MEDYTKIEKI GEGTYGVVYK GRHKTTGQVV AMKKIRLESE EEGVPSTAIR EISLLKELRH 60
 PNIVSLQDVL MQDSRLYLIF EFLSMDLKKY LDSIPPQQYM DSSLVKSYLY QILQGIVFCH 120

SRRLVLRDLK PQNLIDDDKG TIKLADFLGA RAFGIPIRVY THEVVTLLWYR SPEVLLGSSAR 180
 YSTFVDIWSI GTIFAEALTK KPLFHDSEI DQLFRIFRAL GTPNNEVWPE VESLQDYKNT 240
 PPKWKPGSLA SHVNLIDENG LDLLSKMLIY DPAKRISGRM ALNHPYFNDL DNQIKKM

Seq ID NO: 241 DNA sequence

Nucleic Acid Accession #: NM_033379.1

Coding sequence: 132-854

1 11 21 31 41 51
 CGCCGCGCGG CGGGCTCAAC TTGTAGAGC GAGGGGCCAA CTTGGCAGAG CGCGCGGCCA 60
 GCTTTGCAGA GAGCGCCTC CAGGGACTAT GCGTGCGGGG ACACGGGATC TACCCATACC 120
 ATTGACTAAC TATGGAAGAT TATACCAAAA TAGAGAAAAT TGGAGAAGGT ACCTATGGAG 180
 TTGTGTATAA GGGTAGACAC AAAACTACAG GTCAAGTGGT AGCCATGAAA AAAATCAGAC 240
 TAGAAAGTGA AGAGGAAGGG GTTCCTAGTA CTGCAATTCT GGAATTTCT CTATTAAAGG 300
 AACTTGCTCA TCCAAATATA GTCACTCTC AGGATGTGCT TATGCAGGAT TCCAGGTTAT 360
 ATCTCATCTT TGAGTTTCTT TCCATGGATC TGAAGAAATA CTTGGATTCT ATCCCTCCTG 420
 GTCAGTACAT GGATTCTTCA CTGTTAAGG TAGTAACACT CTGGTACAGA TCTCCAGAAG 480
 TATTGCTGGG GTGACGCTCGT TACTCAACTC CAGTTGACAT TTGAGTATA GGCAACCAT 540
 TTGCTGAAC AGCAACTAAG AAACCACTTT TCCATGGGGA TTCAGAAAT GATCAACTCT 600
 TCAGGATTTT CAGAGCTTTG GGCACCTCCA ATAATGAAGT GTGGCCAGAA GTGGAATCTT 660
 TACAGACTTA TAGAATAACA TTTCCCAAT GGAACCAGG AAGCTAGCA TCCCATGTCA 720
 AAAACTTGA TGAAGATGGC TTGGATTGTC TCTGAAAT GTTAATCTAT GATCCAGCCA 780
 AACGAATTTT TGGCAAAATG GCACTGAATC ATCCATATT TAATGATTG GACAATCAGA 840
 TTAAGAAGAT GTAGCTTTCT GACAAAAAGT TTCCATATGT TATGTCAACA GATAGTTGTG 900
 TTTTATTTGT TAACCTTTGT CTATTTTGT CTTATATATA TTTCTTGT ATCAAACTTC 960
 AGCTGTACT CGTCTTCTAA TTTCAAAAT ATAACTTAAA AATGTAAATA TTCTATATGA 1020
 ATTTAAATAT AATTCTGTAA ATGTGAAAAA AAAAAA

Seq ID NO: 242 Protein sequence:

Protein Accession #: NP_203698.1

1 11 21 31 41 51
 MEDYTKIEKI GEGTYGVVYK GRHKTGQVV AMKKIRLESE EEGVPSTAIR EISLLKELRH 60
 PNIVSLQDVL MQDSRLYLIF EFLSMDLKYK LDSIPPQYQM DSSLVKVVTL WYRSPVLLG 120
 SARYSTPVDI WSIGTIFAEI ATKPLFHDG SEIDQLFRIF RALGTNNNEV WPEVESLQDY 180
 KNTFPKWKPG SLASHVKNLD ENGLDLLSKM LIYDPAKRIS GKMLNHPYF NDLDNQIKKM

Seq ID NO: 243 DNA sequence

Nucleic Acid Accession #: AF101051.1

Coding sequence: 221-856

1 11 21 31 41 51
 GAGCAACCTC AGCTTCTAGT ATCCAGACTC CAGGCGCGCC CCGGGCGCGG ACCCCAACCC 60
 CGACCCAGAG CTCTCTCCAG GCGGCGCGAG CGAGCAGGGC TCCCGCCTT AACTTCTCTC 120
 GCGGGGCCCA GCCACCTTCG GGAGTCCGGG TTGCCCACTC GCAAACTCTC CGCTTCTGCG 180
 ACCTGCCACC CCTGAGCCAG CGCGGGCGCC CGAGCGAGTC ATGGCCAAAG CGGGGCTGCA 240
 GCTGTGGGCG TTCAATTCTCG CCTTCTGGG ATGGATCGGC GCCATCGTCA GCACTGCCCT 300
 GCCCAGTGG AGGATTACTC CCTATGCCCG CGACAACATC GTGACCGCCC AGGCCATGTA 360
 CGAGGGGCTG TGGATCTCTC GCGTGTGCGA GAGCACCAGG CAGATCCAGT GCAAGTCTT 420
 TGACTCCTTG CTGAATCTGA GCAGCAGATC CCAAGCAACC CGTGCCTTGA TGGTGGTTGG 480
 CATCTCTCTG GGAGTGATAG CAATCTTTGT GCCCACCCTT GCATGAAGT GTATGAAGTG 540
 CTTGGAAGAC GATGAGGTGC AGAAGATGAG GATGGCTGTC ATTGGGGGTG CGATATTCTC 600
 TCTTGAGGT CTGGCTATTT TAGTTGCCAC AGCATGTGAT GGCAATAGAA TCGTTCAAGA 660
 ATTCTATGAC CCTATGACCC CAGTCAATGC CAGGTACGAA TTTGTTCAGG CTCTCTTCA 720
 TGGCTGGGCT GCTGCTTCTC TCTGCCCTCT GGGAGGTGCC CTACTTTGCT GTTCTGTGCC 780
 CCGAAAAACA ACCTCTTACC CAACACCAAG GCCCTATCCA AAACCTGCAC CTTCCAGCGG 840
 GAAAGACTAC GTGTGACACA GAGGCAAAAG GAGAAAAATCA TGTGAAACA AACCAGAAAT 900
 GGACATTGAG ATACTATCAT TAACATTAGG ACCTTAGAAT TTTGGGTATT GATATCTGAA 960
 GTATGGTATT ACAAAAACA CAACAAACA AAAAAACCAT GTGTAAAAAT ACTCAGTGCT 1020
 AAACATGGCT TAATCTTATT TTATCTTCTT TCCTCAATAT AGGAGGGAAG ATTTTACCAT 1080
 TTGTATTACT GCTTCCCAAT GAGTAATCAT ACTCAATGG GGAAGGGGT GCTCCTTAAA 1140
 TATATATAGA TATGTATATA TACATGTTTT TCTATTAATA ATAGACAGTA AAATACTATT 1200
 CTCATTATGT TGATCTAGC ATACTTAAAA TATCTCTAAA ATAGTAAAT GTATTTAATT 1260
 CCATATTGAT GAAGATGTTT ATGGTATAT TTTCTTTTTC GTCCCTATAT ACATATGTAA 1320
 CAGTCAAAAT TCATTACTC TTCTTCATTA GCTTTGGGTG CCTTTGCCAC AAGACCTAGC 1380
 CTAATTACC AAGATGAAT TCTTTCAATT CTTCTGCGT GCCCTTTTCA TATACTTATT 1440
 TTATTTTTTA CCATAATCTT ATAGCACTTG CATGTTTATT AAGCCCTTAT TTGTTTTGTG 1500
 TTTCATTGGT CTCTATCTCC TGAATCTAAC ACATTTCATA GCCTACATT TAGTTTCTAA 1560
 AGCCAAGAAG AATTATTATC AAATCAGAAC TTTGGAGGCA AATCTTCTG CATGACCAAA 1620
 GTGATAAAT CCTGTGACC TTCCACACA ATCCCTGTAC TGTGCCCAT AGCACTCTTG 1680
 TTTGCTTTGA AAATATTGT CCAATTGAGT AGCTGCATGC TGTTCGCCA GGTGTGTGAA 1740
 CACAACCTTA TTGATTGAAT TTTTAAGCTA CTTATTCTAT GTTTTATATC CCCCTAACT 1800
 ACCTTTTGT TCCCAATCC TTAATTGTAT TGTTTTCCCA AGTGTAATTA TCATGCGTTT 1860
 TATATCTCC TAATAAGGTG TGGTCTGTTT GTCTGAACAA AGTGCTAGAC TTTCTGGAGT 1920
 GATAATCTGG TGACAAATAT TCTCTCTGTA GCTGTAAGCA AGTCACTTAA TCTTCTTACC 1980
 TCTTTTTCT ATCTGCCAAA TTGAGATAAT GATACTTAAC CAGTTAGAAG AGGTAGTGTG 2040
 AATATTAATT AGTTATATT ACTCTCATTC TTTGAACATG AACTATGCCT ATGTAGTGTG 2100
 TTTATTGTG CAGCTGGCTG AGACACTGAA GAAGTCACTG AACAAAACCT ACACAGTATC 2160
 CTTCAATGTA TTCACTGCCT TCCTCTCTCT ACCAGTCTAT TTCCACTGAA CAAAACCTAC 2220
 ACACATACCT TCATGTGGTT CAGTGCCTTC CTCTCTCTAC CAGTCTATT CCACCTGAACA 2280
 AAACCTACGC ACATACCTTC ATGTGGCTCA GTGCTTCTCT CTCTCTACCA GTCTATTPTC 2340
 ATTCTTCTAG CTGTGCTGTA CATGTTTGTG CTCTGTTCCA TTTTAAACAC TGCTCTTACT 2400
 TTTCCAGTCT GTACAGAAAT CTATTTCACT TGAGCAAGAT GATGTATGGA AAGGGTGTG 2460

5 GCACCTGGTGT CTGGAGACCT GGATTGAGT CTTGGTGCTA TCAATCACCG TCTGTGTTTG 2520
AGCAAGGCAT TTGGCTGCTG TAAGCTTATT GCTTCATCTG TAAGCGGTGG TTTGTAAATC 2580
CTGATCTTCC CAOCTCACAG TGATGTTGTG GGGATCCAGT GAGATAGAAT ACATGTAAGT 2640
GTGGTTTGTG AATTGAAAGG GTGCTATACT AAGGGAAGA ATTGAGGAAT TAACTGCATA 2700
CGTTTGGTGG TTGCTTTTCA AATGTTTGA AATAAAAAA TGTTAAGAAA TGGGTTTCTT 2760
GCCTTAACCA GTCTCTCAAG TGATGAGACA GTGAAGTAAA ATTGAGTGCA CTAACGGAAT 2820
AAGATTCTGA GGAAGTCTTA TCCTCTGCAG TGAGTATGGC CCAATGCTTT CTGTGGCTAA 2880
ACAGATGTAA TGGGAAGAAA TAAAGCCTA CGTGTGGTA AATCCAACAG CAAGGGAGAT 2940
TTTTGAATCA TAATAACTCA TAAGTGCTA TCTGTTTCAGT GATGCCCTCA GAGCTCTGTC 3000
10 TGTTAGCTGG CAGCTGACGC TGCTAGGATA GTTAGTTTGG AATGGTACT TCATAATAAA 3060
CTACACAAGG AAGCTCAGCC ACCGTGCTTT ATGAGGAATT GGACCTAATA AATTTTAGTG 3120
TGCCCTCCAA ACCTGAGAAT ATATGCTTTT GGAAGTTTAA ATTTAAATGG CTTTTGCCAC 3180
ATACATAGAT CTTCATGATG TGTGAGTGTA ATTCCATGTG GATATCAGT ACCAAACATT 3240
ACAAAAAAT TTTATGGCC AAAATGACCA ACGAAATGT TACAATAGAA TTTATCCAAT 3300
15 TTTGATCTTT TTATATCTTT CTACCACCC TGGAAACAGA CCAATAGACA TTTTGGGGTT 3360
TTATAATGGG AATTGTGATA AAGCATTACT CTTTTCAT AAATGTTTT TTAATTTAA 3420
AAAAGGAAA AAAAAAAA AAA

20 Seq ID NO: 244 Protein sequence:
Protein Accession #: AAD16433.1

1 11 21 31 41 51
25 MANAGLQLLG FILAFLGWIG AIVSTALPQW RIYSYAGDNI VTAQAMYEGE WMSCVSQSTG 60
QIQCKVFDLS LNLSSLTQAT RALMVVGILL GVIAIPVATV GMKMKCLEED DEVQKMRNAV 120
IGGAIFLLAG LAILVATAWY GNRIVQEFYD PMTPVNARYB FGQALFTGWA AASLCLLGG 180
LLCCSPRKAT TSYTPRPYP KPAFSSGKDY V

30 Seq ID NO: 245 DNA sequence
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51
35 TTTTTTTTTT TTTTTTTTTT TTTTCAAGG AGAGCACAAG GAACCTTATT AATGACTTTC 60
TTAATGGTTA AATGCTGTTT ACCAAGTGAC CCAGAGGCAG CGTGGTTAG TGGTTTCAAC 120
AGCATGGTCC CGAGAGTCTG ACAAACCTCA GTTCAATCC TTCTTTTGTG TCACTTAGT 180
TTTCTTCCCT GAGATTAGT TTCTTCATCG TTAACAATGA GGATATTAAAT ATGTTTCACA 240
CAGTTGTTAT GAAGAATGCA TATATTAGAA TGCCTGTAGT CTCAGTACT CAGGAGGCTA 300
40 AGGTGGGGAG GTGCTCAAG CCCAGGAATT CAAAGCTGCA ATGCATTATG ATTACAGCTG 360
TTAATAGCCA CTGCACTTCA GCCTGGGCAA TGTAGTAAGA TCCCATCTCT GGCTCGGAGG 420
GTCTACGCC CAGGAGTCT CGCTGATTGC TAGCACAGCA GTCTGAGATC AAATGCA

45 Seq ID NO: 246 DNA sequence
Nucleic Acid Accession #: XM_058553.2
Coding sequence: 897-1400

1 11 21 31 41 51
50 AATTTTCAGA AGTTTCGTAT GGGGATGGT TTATATAAAT TCAGGTTTTT CCCACAATAA 60
TAAATGTATT TAGTCTCAGT GCTCAATAGA AGAGATTCTT AATAGAAAAG GATTCAAAC 120
GTGAAACCAT TTCTCTTTTA ATGTTTCACA TTCTGTTTAC AGATTGTGTC TCTTGTGACT 180
CTGTTATCCA TAATATGGAC AGTTCTTGAG TCCTAACATT GAGAGGTTTT CCCTTAGTGC 240
ATAGAGGGAA TGAGTATTAA TTGAGAAGC TTAAGTATT GCCACTTTAG CACTGAAGAT 300
55 TGGGATGAGA GAGGTGAA GCTCACTAGA AAAAGGGACA ATGTTAGTGT GGCCCTTCTC 360
GATCATGTTT AAGAAAAGTC ATGAAAATGG TGAAGTAGTG TTTCCAAGCA TATTGGAAGG 420
GTTGAGTGTA TACTGTCTGT CAAAGACTTC CAGCATTTCC AGGTCCTAGA GAGGAACAAG 480
ACTGGTAACC TGCTATCTG TATTTTAAAG AACCCAGGAG GAAAGCTTTA TAATAGAACA 540
TTATTTCTGT GTTTATGTAT AAGGGGTTTT TTGTTTTTTT AAAGACAGGA TCTCACTCCA 600
60 TTGTCCAGGC CAAGTGCAAT GGCACGAACC TCATAGCTCC TGGACTTAAG TGATCTGCCT 660
GCCTTTGCCCT CCTGAGTAGC TGGGACTACA GGCATGAGCC CCCATGCCCT GCTAAGTTTG 720
TTTTTTTGTG TTGTTGTTTG TTTGTTTTTG GGGGGGGTTG TTTTGTTTTT TGTAGAGACG 780
TAGTCTGCT TTGTTGCCAG GCTAGTCTCA AACTCCTGGC TTCAGTGAT CCTCTGCTCT 840
CAGCCTCCCA GAGTGTAGG ATTACAGCAC TTGGATTTCAG CTCTCTTCAAT TCCAACATGG 900
65 AAGAAACTTA CACGACTCC CTGGACCTG AGAAGCTATT GCAATGCCCC TATGACAAAA 960
ACCATCAAAAT CAGGGCTTGC AGGTTTCCTT ATCATCTTAT CAAGTCGACA AAGAATCATC 1020
CTGATGTGCG AAGCAAAATG GCTACTGTGC CCTTCAATGC TCGCCACCAG GTTCTCTCAG 1080
CTGAAATTAG TCATCATATC TCAAGCTGTG ATGACAGAAG TTGTATTGAG CAAGATGTTG 1140
TCAACCAAAAC CAGGAGCCTT AGACAAGAGA CTCTGGCTGA GAGCACTTGG CAGTGCCTCT 1200
70 CTTCGATGA AGACTGGGAT AAAGATTGTG GGGAGCAGAC CAGCACCCCA TTTGTCTGGG 1260
GCACAACCTCA CTACTCTGAC AACACAGGCC CTGCGAGCAA CATAGTTACA GAACATAAGA 1320
ATAACCTGGC TTCAGGATG CGAGTTCCCA AATCTCTGCC GTATGTTCTG CCATGGAAAA 1380
ACAAATGAAA TGCACAGTAA CTGAATACCT ATCTCATCAA ATGCCAGACC CTAGAAGACT 1440
75 GTTGCTTCTT CTCTACCAAG TGGGTTCTCA TTTTCTCTCT AATCTAATTA TAGAATGGTA 1500
AACTCCCTGT GACTTTCCAA ACTGACAAGC ACCTTTTCTT CCTCCCCCT TGAATCCTCA 1560
TTTAATGCAA GAACCTCAT ACTCAGAAGC TTCCAATAA ACCTTTGATA CAGATTG

80 Seq ID NO: 247 Protein sequence:
Protein Accession #: XP_058553.1

1 11 21 31 41 51
85 MEETYSLSLD PEKLLQCPYD KNHQRACRF PYHLIKCRKN HPDVASKLAT CPFNARHQVP 60
RAEISHHISS CDRSCEIQD VVNQTRSLRQ STLAESTWQC PPCDEWDKRD LWEQTSTPFV 120
WGTHYSDNN SPASNVITEH KNINLASGMRV PKSLPYVLPW KNNGNAQ

WO 02/086443

PCT/US02/12476

Seq ID NO: 248 DNA sequence
Nucleic Acid Accession #: NM_003392
Coding sequence: 758..1855

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
TTAAGGAAT CCGGGCTGCT CTGCCCATC TGGAGGTGGC TTTCCCCACA TGGCTCGTA 60
AACTGATTAT GAAACATACG ATGTTAATTC GGAGCTGCAT TTCCCAGCTG GGCACTCTCG 120
CGCGCTGGTC CCGGGGGCTT CGCCCCCACC CCCCTGCCCT TCCCTCCCGC GTCCCTGCCCC 180
CATCTCCAC CCCCCGGGCT GGGCCACCCG CCTCCTTGCG AGCCTCTGGC GGCAGGGGCG 240
TCCACTCGCC TCCCGTGCTC CTCTGCCCCA TGGAAATTAAT TCTGGCTCCA CTGTGTGCTC 300
GGCCAGGTT GGGGAGAGGA CGGAGGGTGG CCGCAGCGGG TTCTGTAGTG AATTACCCAG 360
GAGGGACTGA GCACAGCACC AACTAGAGAG GGGTCAGGGG GTGCGGGAAT CGAGCGAGCA 420
GGAAGGAGGC AGCGCTGGC ACCAGGGCTT TGACTCAACA GAATTGAGAC ACGTTTGTAA 480
TGCTGGGCT GCGCCGCGCA CAGGATCCCA GCGAAATCA GATTCTCTGG TGAGGTTGCG 540
TGGGTGATTA AATTGGGAAA AAGAAACTGC CTATATCTTG CCATCAAAA ACTCACGGAG 600
GAGAAGCGCA GTCAATCAAC AGTAACTTA AGAGACCCCC GATGCTCCCC TGGTTAACT 660
TGTAATCTTG AAAATATCTT GAGAGGGAAT AAACATCTTT TCCTTCTTCC CTCTCCAGAA 720
GTCCATTGGA ATATTAAGCC CAGGAGTTGC TTTGGGGATG GCTGGAAGTG CAATGTCTTC 780
CAAGTTCTTC CTAGTGGCTT TGGCCATATT TTTCTCCTTC GCCCAGTTTG TAATTGAAGC 840
CAATTCTTGG TGGTGGCTAG GTATGAATAA CCTGTTCAG ATGTCAAGAG TATATATTAT 900
AGGAGCACAG CTCCTCTGCA GCCAATCGGC AGGACTTTCT CAAGGACAGA AGAACTGTG 960
CCACTTGTAT CAGGACCAAC TGCACTACAT CGGAGAGGCG GCGAAGACAG ACATCAAGA 1020
ATGCCAGTAT CAATTCCGAC ATCGACGGTG GAACCTGCAG ACTGTGATA ACACCTCTGT 1080
TTTTGGCAGG GTGATGCGAG TAGGACGGCG CGAGACGGCC TTCACATACG CCGTGAGGCG 1140
AGCAGGGGTG GTGAACGCCA TAGCCGGGCG GTGCGCGGAG GCGGAGCTGT CCACCTGCGG 1200
CTGCAGCGCG CCGCGCGGCC CCAAGGACCT GCCCGGGGAC TGGCTCTGGG GCGGCTGCGG 1260
CGACACATC GACTATGGCT ACCGCTTTGC CAAGGAGTTC GTGGAAGCCC GCGAGCGGGA 1320
GCGCATCCAC GCCAAGGGCT CCTACGAGAG TGCTCGCATC CTCTAGAACG TGCAACAA 1380
CGAGGCGCGC CGCAGGACGG TGTAACAACCT GGCTGATGTG GCCTGCAAGT GCCATGGGT 1440
GTCCGGCTCA TGAGCCTTGA AGACATGCTG GCTGCAGCTG CAGAGCTTCC GCAAGGTGGG 1500
TGATGCCCTG AAGGAGAAGT ACGACAGCGC GCGCGGCCAT GCGCTCAACA GCGGGGCAA 1560
GTTGGTACAG GTCAACAGCC GCTTCAACTC GCCCACCACA CAAGACCTGG TCTACATCG 1620
CCCCAGCCTT GACTACTGCG TGCGCAATGA GAGCACCGGC TCGCTGGGCA CGCAGGGCG 1680
CCTGTGCAAC AAGAGCTCGG AGGGCATGGA TGGCTGCGAG CTCACTGTCT GCGGCGGTG 1740
GTAAGACAG TTCAAGACCG TGCAAGCGGA GCGCTGCCAC TGCAAGTTCC ACTGGTGTG 1800
CTACGTCAAG TGCAAGAAGT GCACGGAGAT CGTGGACCAG TTTGTGTGCA AGTAGTGGGT 1860
GCCACCCAG ACTCAGCCCC GCTCCAGGAA CCGCTTATT TATAGAAAGT ACAGTGATT 1920
TGGTTTTTGG TTTTATGAAA TATTTTTTAT TTTTCCCCAA GAATTGCAAC CGGAACCAAT 1980
TTTTCCTCTT TACCATCTA AGAACTCTGT GGTATTATTA TAATATTATA ATTATTATT 2040
GGCAATAATG GGGGTGGGAA CCACGAAAAA TATTTATTTT GTGGATCTTT GAAAGGTAA 2100
TACAAGACTT CTTTGGGATA GTATAGAATG AAGGGGAAA TAACACATAC CCTAACTTAG 2160
CTGTGTGGGA CATGGTACAC ATCCAGAAGG TAAGAAATAA CATTTCTTTT TCTCAATA 2220
TGCCATCAT TGGGATGGGT AGGTTCCAGT TGAAAGAGGG TGGTAGAAAT CTATTACAA 2280
TTCACTCTCT ATGACCAAAA TGAGTTGTAA ATCTCTGTGT GCAAGATAAA AGGCTTGGG 2340
AAAACAAAA AAAACAAAAC AAACCTCCCT TCCCAGCAG GGTCTGTAGC TTGCTTTCTG 2400
CATTTTCAAA ATGATTAATT ACAATGGAAG GACAAGAAAT TCATATTCTC AAGGAAAAAA 2460
GGTATATCAC ATGTCTCATT CTCCTCAAA ATTCCATTG CAGACAGACC GTCATATTCT 2520
AATAGCTCAT GAAATTTGGG CAGCAGGGAG GAAAGTCCCC AGAAATTAAG AAATTTAAAA 2580
CTCTTATGTC AAGATGTGGA TTTGAAGCTG TTATAAGAA TGGGATTCCA GATTGTGAAA 2640
AAGACCCCCA ATGATTCTGG ACACAGATT TTTGTGTTGG GGAGGTGGC TTGAACATAA 2700
ATGAAATATC CTGATTTTTC TTAGGGATAC TTGGTTAGTA AATTATAATA GTAGAAATAA 2760
TACATGAATC CCATTACAG GTTTCTCAGC CCAAGCAACA AGGTAATTGC GTGCCATTCA 2820
GCACTGACC AGAGCAGACA ACCTATTGGA GGAAACACAG TGAAATCCAC CTTCCTCTTC 2880
ACACTGAGCC CTCTCTGATT CCTCCGTGTT GTGATGTGAT GCTGGCCACG TTTCCAAAAG 2940
GCAGCTCCAC TGGGTCCCTT TTGGTTGTAG GACAGGAAAT GAAACATTAG GAGCTCTGCT 3000
TGGAAACAG TTCACTACTT AGGGATTTTT GTTTCCTAAA ACTTTTATT TTGAGGAGCAG 3060
TAGTTTTCTA TGTTTTAATG ACAGAACTTG GCTAATGGAA TTCACAGAGG TGTGCGAGCG 3120
TATCACTGTT ATGATCTGTT GTTTAGATTA TCCACTCATG CTCTCTCTAT TGTACTGCG 3180
GTGTACTCTA AAAGTGTCC CAGTGTACTT GAACAGTTGC ATTTATAAGG GGGGAAATGT 3240
GGTTTAATGG TGCCGTATAT CTCAAAGTCT TTTGTACATA ACATATATAT ATATATACAT 3300
ATATATAAAT ATAAATATAA ATATATCTCA TTGCAGCCAG TGATTAGAT TTACAGCTTA 3360
CTCTGGGGTT ATCTCTCTGT CTAGAGCATT GTTGCTCTTC ACTGCAGTCC AGTTGGGATT 3420
ATTCCAAAG TTTTGTGAGT CTTGAGCTTG GGCTGTGGCC CCGCTGTGAT CATACCTTGA 3480
GCACGACGAA GCAACCTGCT TTCTGAGGAA GAAGCTGTAG TTCTGACTCA CTGAAATGCG 3540
TGTGCGGTG AAGATATCTT TTTTCTTTT CTGCTCACC CCTTTGTCTC CAACCTCCAT 3600
TTCTGTTCAC TTTTGGGAGA GGGCATTACT TGTTCGTTAT AGACATGAC GTTAAGAGAT 3660
ATTCAAACT CAGAAGCATC AGCAATGTTT CTCTTTTCTT AGTTCACTCT CGCAATGGA 3720
AACCATGCC TATTAGAAAT GACAGTACTT ATTAATTGAG TCCCTAAGGA ATATTGAGC 3780
CACTACATAG ATAGCTTTT TTTTCTTTT TTTTCTTTT TAAGGACACC TCTTTCCAAA 3840
CAGGCCATCA AATATGTTCT TATCTCAGAC TTAAGTTGTT TTAAGTTT TGAAGATAC 3900
ACATCTTTTC ATACCCCTCC TTAGGAGGT GGGCTTTCT ATCACCTCAG CCACTGTGG 3960
CTCTTAATTT ATTGCATAAT GATATCCACA TCAGCCAACT GTGGCTCTTT AATTATTGCG 4020
ATAATGATAT TCAATCTAG TCAATGTGGA GCAAAAGATC TGAAGAGCAA 4080
AAAGCACTAA TTAGTTTAAA ATGTCACTTT TTTGGTTTT ATTATACAAA AACCATGAAG 4140
TACTTTTTTT ATTTGCTAAA TCAGATTGTT CCTTTTGTAG GACTCATGTT TATGAAGAGA 4200
GTTGAGTTTA ACAATCTAG CTTTTAAAAG AAACATTTTA ATGTAATAA TTCTACATGT 4260
CATTCAGATA TTAGTATAT CTCTAGCCT TATTTCTGTA CTTTTAAATG ACATATTCT 4320
GTCTGCGTG ATTTGATAT TTCACTGGTT TAAAAACAA ACATCGAAG GCTTATTCCA 4380
AATGGAAGAT AGAATATAA ATAAACGTT ACTTGTAATA AAAAAAAA

Seq ID NO: 249 Protein sequence:
Protein Accession #: NP_003383

1 11 21 31 41 51
| | | | | |

MAGSAMSSKF FLVALAIFFS FAQVVEANS NWSLGMNPFV QMSEVYIIGA QPLCSQLAGL 60
 SQGQKLLCHL YQDHMQYIGE GAKTGIKECQ YQFRHRRHNC STVDNTSVFG RVMQIGSRET 120
 APTYAVSAAG VVNAMSRACR EGELSTCGCS RAARPDLPDR DMLWGGCGDN IDYGYRFAKE 180
 FVDARERERI HAKGSYESAR ILMNLEENEA GRRTVYNLAD VACKCHGVSG SCSLKTCWLQ 240
 LADPRKVGDA LKEKYDSAAA MRLNSRGKLV QVNSRFNSPT TQDLVYIDPS PDCVNRNEST 300
 GSLGTQGRLC NKTSEGMDCG ELMCCGRGYD QFKTVQTERC HCKFWCCYV KCKKCTEIVD 360
 QFVCK

Seq ID NO: 250 DNA sequence
 Nucleic Acid Accession #: NM_014058
 Coding sequence: 56..1324

1 11 21 31 41 51
 TGACTTGGAT GTAGACCTCG ACCTTCACAG GACTCTTCAT TGCTGGTGG CAATGATGTA 60
 TCGGCCAGAT GTGGTGGGG CTAGGAAAAG AGTTTGTGG GAACCTGGG TTATCGGCT 120
 CGTCACTTTC ATATCCCTGA TTGTCTGGC AGTGTGCATT GGACTCAGTG TTCATTATGT 180
 GAGATATAAT CAAAAGAAGA CCTACAATTA CTATAGCACA TTGTCAATTA CAACTGACAA 240
 ACTATATGCT GAGTTTGGCA GAGAGGCTTC TAACAATTTT ACAGAAATGA GCCAGAGACT 300
 TGAATCAATG GTGAAAATG CATTTTATAA ATCTCCATTA AGGGAAGAAT TTGTCAAGTC 360
 TCAGGTTATC AAGTTCAGTC AACAGAAGCA TGGAGTGTG GCTCATATGC TGTGATTG 420
 TAGATTTTAC TCTACTGAGG ATCCTGAAAC TGTAGATAAA ATTGTTCAAC TTGTTTACA 480
 TGAAAAGCTG CAAGATGCTG TAGGACCCCG TAAAGTAGAT CCTCACTCAG TTAATAATTA 540
 AAAAACTAAC AAGACAGAAA CAGACAGCTA TCTAAACCAT TGCTGCGGAA CACGAAGAAG 600
 TAAACTCTGA GGTCAAGATG TCAGGATCGT TGGTGGGACA GAAGTAGAAG AGGGTGAATG 660
 GCCCTGGCAG GCTAGCCTGC AGTGGGATGG GAGTCATGCG TGTGAGCAA CCTTAATTA 720
 TGCCACATGG CTGTGTGAGTG CTGCTCACTG TTTTACAACA TATAAGAACC CTGCCAGATG 780
 GACTGCTTCC TTTGGAGTAA CAATAAAACC TTGAAAATG AAACGGGTC TCCGGAGAAT 840
 AATTGTCCAT GAAAATACA AACACCCATC ACATGACTAT GATATTTCTC TTGCAGAGCT 900
 TTCTAGCCCT GTTCCCTACA CAAATGCAGT ACATAGAGTT TGCTCCCTG ATGCATCTTA 960
 TGAGTTTCAA CCAGGTGATG TGATGTTGT GACAGGATTT GGAGCACTGA AAAATGATGG 1020
 TTACAGTCAA AATCATCTTC GACAAGCACA GGTGACTCTC ATAGACGCTA CAACTTGCAA 1080
 TGAACCTCAA GCTTACAATG ACGCCATAAC TCCTAGAATG TTATGTGCTG GCTCCTTAGA 1140
 AGGAAAAACA GATGATGCC AGGGTGACTC TGGAGGACCA CTGGTTAGTT CAGATGTCTAG 1200
 AGATATCTGG TACCTTGTCTG GAATAGTGAG CTGGGAGAT GAATGTGGA AACCCAACA 1260
 GCCTGGTGT TATACAGAG TTAGCGCCTT GCGGACTGG ATTACTTCAA AAAGTGGTAT 1320
 CTAAGAGAGA AAAGCCTCAT GGAACAGATA ACATTTTTTT TTGTTTTTTG GGTGTGGAGG 1380
 CCATTTTTAG AGATACAGAA TTGGAGAAGA CTTCGAAAC AGCTAGATT GACTGATCTC 1440
 AATAAATCTG TTGCTTGATG CAAAAAATA A

Seq ID NO: 251 Protein sequence:
 Protein Accession #: NP_054777

1 11 21 31 41 51
 MYRPDVVRAR KRVCWEPWVI GLVIFISLIV LAVCIGLTVH YVRYNQKITY NYYSTLSFTT 60
 DKLYAEFGRE ASNNFTFMSQ RLESMVRNAF YKSPLEEFV KQVQIKFSQ KQGVLAHMLL 120
 ICRPHSTEDP ETVDKIVQLV LHEKLQDQAG PPKVDPHSVK IKKINKTETD SYLNHCOGTR 180
 RSKTLGSLR IVGGTEVEEG EWPWQASLQW DGSHRGATD INATWLVSAA HCFPTYKNPA 240
 RWTASFGVTI KPSKMKRGLR RIIVHEKYKH PSHDYDISLA ELSPPVPTN AVHRVCLPDA 300
 SYEFQPGDVM FVTGFGALKN DGYSQNHRLR AQVTLIDATT CNEPQAYNDA ITPRMLCAGS 360
 LBGKTDACQG DSGGPLVSSD ARDIWYLAGI VSWGDECAKP NKPGVYTRVT ALRDWITSKT 420
 GI

Seq ID NO: 252 DNA sequence
 Nucleic Acid Accession #: NM_003504.2
 Coding sequence: 71-1771

1 11 21 31 41 51
 GGCACGAGGC CTGCTGCCGC CGGGCTCTTG GTACCTCAGC GCGAGGCGCA GGGGTCCGGC 60
 CGCCGTGGCT ATGTTGCTGT CCGATTTCGG CAAAGAGTTC TACGAGGTGG TCCAGAGCCA 120
 GAGGGTCTCT CTCTTCGTGG CCTCGGACGT GGATGCTCTG TGTGGTGCA AGATCCTTCA 180
 GGCTTGTGTC CAGTGTGACC ACGTGCAATA TACGCTGGTT CCAGTTTCTG GGTGGCAAGA 240
 ACTTGAACAT GCATTCTCTG AGCATAAAGA ACAGTTTCAT TATTTTATTC TCATAAACCTG 300
 TGGAGCTAAT GTAGACCTAT TGGATATTCT TCAACTGAT GAAGACACTA TATCTTTGT 360
 GTGTGACACC CATAGGCCAG TCAATGCTGT CAATGTATAC AACGATACCC AGATCAAATT 420
 ACTCATTAAG CAAGATGATG ACCTTGAAGT TCCCGCTAT GAAGACATCT TCAGGATGTA 480
 AGAGGAGGAT GAAGAGCATT CAGGAAATGA CAGTGATGGG TCAGAGCCTT CTGAGAAGCG 540
 CACACGTTTA GAAGAGGAGA TAGTGGAGCA AACCATGCGG AGGAGGCAGC GGCAGAGATG 600
 GGAGGCGCGG AGAAGAGACA TCCTCTTTGA CTACGAGCAG TATGAATATC ATGGGACATC 660
 GTCAGCCATG GTGATGTTTG AGCTGGCTTG GATGCTGTCC AAGGACCTGA ATGACATGCT 720
 GTGGTGGGCC ATCGTTGGAC TAACAGACCA GTGGGTGCAA GACAAGATCA CTCAAATGAA 780
 ATACGTGACT GATGTTGGTG TCCTGACGCG CCACGTTTCC CGCCACAACC ACCGGAACGA 840
 GGATGAGGAG AACACACTCT CGGTGACTG CACACGGATC TCCTTTGAGT ATGACCTCCG 900
 CCTGGTGTCT TACCAGCACT GGTCCCTCCA TGACAGCCTG TGCAACACCA GCTATACGGC 960
 AGCCAGGTTT AAGCTGTGGT CTGTGATGG ACAGAAGCGG CTCGAGGAGT TCCTTGACGA 1020
 CATGGGTCTT CCCTCGAAGC AGGTGAAGCA GAAGTTCCAG GCCATGGACA TCTCTTGAA 1080
 GGAGAAATTG CGGGAATGA TTGAAGAGTC TGCAATAAAA TTGGGATGA AGGACATGCG 1140
 CGTGCACTAT TTCAGCATTC ATTTTGGGTT CAAGCAAGAT TTCTGGCCA GCGACGTGGT 1200
 CTTTGCCACC ATGTCTTTGA TGGAGAGCCC CGAGAAGGAT GGTCTAGGGA CAGATCACTT 1260
 CATCCAGGCT CTGGACAGCC TCTCCAGGAG TACCTGGAC AAGCTGTACC ATGGCCTGGA 1320
 ACTCGCCAG AAGCAGCTGC GAGCCACCCA GCAGACCAAT GCCAGCTGCC TTTCACCAA 1380
 CCTGTCTAT TCCAGGGGCG CATCCCTAAG CCTGCTCAGC AACACCTGC TCAAGTCCTT 1440
 CATGCTGTTC TCTAGGCCGG CATCCCTAAG CCTGCTCAGC AACACCTGC TCAAGTCCTT 1500
 TGTGTGTCG ACAAAGAACC GCGCTGCAA ACTGCTGCCC CTGGTGTATG CTGCCCCCTT 1560

5

GAGCATGGAG CATGGCACAG TGACCGTGGT GGGCATCCCC CCAGAGACCG ACAGCTCGGA 1620
 CAGGAAGAAC TTTTITGGGA GGGCGTTTGA GAAGGCAGCG GAAAGCACCA GCTCCCGGAT 1680
 GCTGCACAC CATTTTGACC TCTCAGTAAT TGAGCTGAAA GCTGAGGATC GGAGCAAGTT 1740
 TCTGGACGCA CTATTATCCC TCCTGTCCTA GGAATTGTAT TCTTCCAGAA TGACCTTCTT 1800
 ATTTATGTAA CTGGCTTTCA TTTAGATTGT AAGTTATGGA CATGATTTGA GATGTAGAAG 1860
 CCATTTTTTA TTAATAAAAA TGCTTATTTT AGGCTCCGTC CCCAAAAAAA AAAAAAAA 1920
 AAAAAAAA AA

10

Seq ID NO: 253 Protein sequence:
 Protein Accession #: NP_003495.1

15

20

25

1 11 21 31 41 51
 MPVSDRKEF YEYVQSRVL LFVSDVDAL CACKILQALF QCDHVQYTLV FVSGWQELST 60
 AFLEHKEQFH YFILINGAN VDLLDILQPD EDTIFFVCDT HRPVNVVNVY NDTQIKLLIK 120
 QDDLEVPAY EDIFRDEED EEHSGNDSG SEPSEKTRRL EEEIVEQTMQ RRQRREWEAR 180
 RRDILFDYEQ YEYHGTSAM VMFELANMLS KDLDMLNWA IVGLTDQWVO DKITQMKYVT 240
 DVGVLQRHVS RHNHRNEDEE NTLSDVCTRI SFEYDLRLVL YQHSLEDSL CNTSYTAARF 300
 KLWSVHGQKR LQEFLLDMGL PLKQVQKFPQ AMDISLKENL REMIEESANK FGMDMRVQT 360
 PSIHFGPKK FLASDVVFAT MSLMESPEKD GSGTDHFIAQ LDSLSRSNLD KLYHGLELAK 420
 KQLRATQQT ASCLCNTLVI SQGFPLYCSL MEGTPDVMLF SRPASLSLLS KHLKSEFVCS 480
 TKNRRCKLLP LVMAAPLSME HGTVTVVGIP PETDSSDRKN PFGRAFEKAA ESTSSRMLHN 540
 HFDLSVIELK AEDRSKFLDA LISLLS

Seq ID NO: 254 DNA sequence
 Nucleic Acid Accession #: NM_022337
 Coding sequence: 48..683

30

35

40

45

50

55

1 11 21 31 41 51
 GGCTGCGCTT CCCTGGTCAG GCACGGCAGC TCTGGCCGCG CGCCAGGATG CAGGCCCGCG 60
 ACAAGGAGCA CCGTGTACAG TTGCTGGTGA TTGGCGACCT GGGCGTGGGG AAGACCAGTA 120
 TCATCAAGCG CTACGTGCAC CAGAACTTCT CCTCGCACTA CCGGGCCACA ATCGGGCTGG 180
 ACTTCGCGCT CAAGGTGCTC CACTGGGACC CGGAGACTGT GGTGCGCCTG CAGCTCTGGG 240
 ATATCGCAGG TCAAGAAAGA TTTGGAAACA TGACGAGGGT CTATTACCGA GAAGCTATGG 300
 GTGCATTTAT TGCTTTCGAT GTCACCAAGC CAGCCACATT TGAAGCAGTG GCAAAGTGGA 360
 AAAATGATTT GGACTCCAAG TTAAGTCTCC CTAATGGCAA ACCGCTTTCG GTGGTTTGTG 420
 TGGCCACAAA ATGTGACCAAG GGGGAAGGATG TGCTCATGAA CAATGGCCTC AAGATGGACC 480
 AGTTCTGCAG GGAGCACGGT TTCGTAGGAT GGTTTGAAAC ATCAGCAAAG GAAAATATAA 540
 ACATTGATGA AGCTCCAGA TGCCTGGTGA AACACATACT TGCAATGAG TGTGACCTAA 600
 TGGAGTCTAT TGAGCCGAGC GTCGTGAAGC CCCATCTCAC ATCAACCAAG GTTGCCAGCT 660
 GCTCTGGCTG TGCCAAATCC TAGTAGGCAC CTTTGTCTGT GTCTGTAGG AATGACCTCA 720
 TTGTTCCACA AATGTGCGCT CTATTTTATC CATTTTGGGT AAACGTGAGG ATAGATATAC 780
 CACATGTGCG AAGCCAAAGA TCTATGCCCTC TGTTTTTTCG ATGAGAGAGA AATAGCAAAT 840
 GTTCTTTCTA TGCTTTCTCT ACCATCATCA CAGTGTTTAC AAACCTTTTG AATATTTTAG 900
 TCTGTTACAA ACTTCTGTCA TGTAGCTGAC CAAAATCCTG CAGGGCCACA GTGGCACTG 960
 TTATTTGCTT CTTTTAATCA GCAAAGGCGT CAAGTCTTAA AATAAAGGG GAGAAGAACA 1020
 AACTAGCTGT CAAGTCAAGG ACTGGCTTTC ACCTTGCCCT GGTGCTTTT TCCAGATTTT 1080
 AATATATCTT CTGATGAGG GACAGGCCTA TTAAGTAGAT GTGATATTTT CTTCCAAGAT 1140
 GAOCTCCATT CTGGGCAGAC CTAAGAGTTG CCTCTGAGT AGCTCTTTGG AATCGTGAAC 1200
 ACAGGTGTGC TATATTGTCC TTGTCCTAAC TGTCACTTGC CATGGCCTGA ATGTTGGCTT 1260
 AACTGAATAT TGTATGAAAA GACATGCCCT CATATGTGCC TTTCTGTTAG CTCTCTTTGA 1320
 CTCAGCTGTG GGGGCTCCTC TATACATGCT ATACATGTAA TATATATTAT ATATATTTTT 1380
 GCAAGTGAAC AATAAACAT TAAAGATAA AA

Seq ID NO: 255 Protein sequence:
 Protein Accession #: NP_071732

60

65

1 11 21 31 41 51
 MQAPHKEHLY KLLVIGDLGV GKTSIIKRVY HQNFSSHYRA TIGVDFALKV LHWDPETVVR 60
 LQLWDIAGQE RFGNMTRVYV REAMGAFIVF DVTRPATFEA VAKWKNDLDS KLSLPNGKPV 120
 SVVLLANKCD QGKDVLMNNG LKMDQFCKEH GFVGVWFETSA KENINIDEAS RCLVKHILAN 180
 ECDLMESIEP DVVKPHLTST KVASCSCGCAK S

Seq ID NO: 256 DNA sequence
 Nucleic Acid Accession #: NM_016321
 Coding sequence: 25..1464

70

75

80

85

1 11 21 31 41 51
 GGAACGCGCC GCTGCCAGCC CGGCCAGGCA CCCCTGCAGC ATGGCCTGGA ACACCAACCT 60
 CCGCTGGCGG CTGCGGCTCA CCGGCTGCT CCTGCAGGTG ATTATGTGTA TTCTCTTCGG 120
 GGTGTTCTGT CGCTACGACT TCGAGGCCGA CGCCCACTGG TGGTCAGAGA GGACGCACAA 180
 GAACCTTGAG GACATGGAGA ACGAATTCTA CTATCGCTAC CCAAGCTTCC AGGACGTGCA 240
 CGTGATGGTC TTGCTGGGCT TCGGCTTCCT CATGACTTTC CTGCAGCGCT ACGGCTTCAG 300
 CGCGCTGGGC TTCAACTTCC TGTGGGAGC CTTGCGCATC CAGTGGGCGC TGCTCATGCA 360
 GGGCTGGTTC CACTTCTTAC AAGACGCTA CATCGTGTG GCGGTGGAGA ACCTCATCAA 420
 CGCTGACTTC TCGGTGGGCT CTGTCTGCGT GGCCTTTGGG GCAGTTCCTG GTAAAGTCAG 480
 CCCCATTCAG CTGCTCATCA TGACTTCTT CCAAGTGACC CTCTTCGCTG TGAATGAGTT 540
 CATTCCTCTT AACCTGCTAA AGGTGAAGGA TGCAGGAGGC TCCATGACCA TCCACACATT 600
 TGGGCGCTAC TTGGGGCTCA CAGTGACCCG GATCCTCTAC CGACGCAACC TAGAGCAGAG 660
 CAAGGAGAGA CAGAAATCTG TGTACCACTG GGACCTCTTT GCCATGATTG GCACCTCTTT 720
 CCTGTGGATG TACTGGCCCA GCTTCAACTC AGCCATATCC TACCATGGGG ACAGCCAGCA 780
 CCGAGCGGCC ATCAACACCT ACTGCTCCTT GGCAGCCTGC GTGCTTACCT CGGTGGCAAT 840

ATCCAGTGGC CTGCACAAGA AGGGCAAGCT GGACATGGT CACATCCAGA ATGCCAGGCT 900
 CGCAGGAGGG GTGGCCGTGG GTACCGCTGC TGAGATGATG CTCATGCCCTT ACGGTGCCCT 960
 CATCATCGGC TTGCTCTGGG GCATCATCTC CACCTGGGT TTGTATACCC TGACCCCAAT 1020
 CCTGGAGTCC CGGCTGCACA TCCAGGACAC ATGTGGCAAT AACATCTGC ATGGCAITCC 1080
 TGGCATCATA GGGCGCATCG TGGGTGCTGT GACAGCGGCC TCGGCCAGCC TTGAAGTCTA 1140
 TGGAAAACAA GGGCTTGTC ATTCCTTTGA CTTTCAAGGT TTCAAAGGGG ACTGGACCGC 1200
 AAGAACACAG GGAAGTTCC AGATTATGG TCTCTTGGT ACCCTGGCCA TGGCCCTGAT 1260
 GGGTGGCATC ATTGTGGGGC TCATTTTGAG ATTACCATT TGGGGACAAC CTTGAGATGA 1320
 GAATGCTTTT GAGGATGCGG TCTACTGGGA GATGCCCTGA GGAACAGCA CTGCTACAT 1380
 CCTGAGGAC CCCACCTTCA AGCCCTCAGG ACCCTCAGTA CCTCAGTAC CCATGGTGTC 1440
 CCCACTACCC ATGGCTTCTT CGGTACCCCTT GGTACCCCTAG GCTCCAGGG CAGGTGAGGA 1500
 GCAGGCTCCA CAGACTTCC TGGGGCCAG AGGAGCTGGT GCTGACCTAG CTAGGGATGC 1560
 AAGAGTGAGC AAGCAGCACC CCCACCTGCT GGCTTGGCT CAAGGTGCTT CCACCCCTGC 1620
 CCTCCCTTTC ATCCAGGGG GTCTGMCCTGA GAATGGAGAA GGAGAAGCTA CAAAGTGGGC 1680
 ATCCAAAGCG GTTCTGGCT GCAGAACTTC TGCCTCTGCC TGGGGTCTTG GCCCATTTGG 1740
 AGAAAAACAG GCTCAAAGTG GGGCTGGGAC CTGGTGGGTG AACCTGAGCT CTCCAGGAG 1800
 ACAACTTAGC TGGCAGTAC CACCTATGAG GCTCTTCTAC CCGCTGCTG CACCTCGGCC 1860
 AGCATCTCCT ATGCTCCCTG GGTCCCCCAG ACCTCTCTGT GTTGTGTGG TGGCAGCCTC 1920
 CAGGAATAAA CATTCTTGTG GTCCTTTGTA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 257 Protein sequence:
 Protein Accession #: NP_057405

1 11 21 31 41 51
 MAWNTNLNRW LPLTCLLLQV IMVILFVGVF RYDFEADAHW WSEKTHKNLS DMENEPYRY 60
 PSQDVHVMV FVGFGLMTF LQRYGPSAVG FNFLAAGFI QWALLMQGW FFLQDRIYV 120
 GVENLINADF CVASVCVAFG AVLKVSPIQ LLIMTFQVLT LFAVNEFIL NLLKVKDAGG 180
 SMTIHTFGAY FGLTVTRILY RRNLQSKER QNSVYQSDLF AMIGTLPLWM YWPSFNGAIS 240
 YHGDSSHRAA INTYCSLAAC VLTSVAISSA LHKKGKLDLV HIQNATLGG VAVGTAAEMM 300
 LMPYGLIIG FVCGIISTLG FVYLTPFLES RLHIQDTGCI MNLHGIPGII GGIVGAVTAA 360
 SASLEVYAGC GLVHSFDFQG FNGDWTARTQ GKFIYGLLV TLAMALMGII IVGLILRLPF 420
 WQPSDENCF EDAVYEMPE GNSTVIYPED PTFKPSGPSV PSVPMVSLP MASSVPLVP

Seq ID NO: 258 DNA sequence
 Nucleic Acid Accession #: NM_002358.2
 Coding sequence: 75..692

1 11 21 31 41 51
 GGGAAAGTGT GTTGGAGCGC CTGTGGTTGC TGTCCGCGGA GTGGAAGCGC GTGCTTTTGT 60
 TTGTGTCCCT GGCCATGGCG CTGCAGCTCT CCGCGGAGCA GGAATCACC CTGCGCGGGA 120
 GCGCGAATAT CTGGGCGGAG TTCTTCTCAT TCGGCATCAA CAGCAITTTA TATCAGCGTG 180
 GCATATATCC ATCTGAAACC TTTACTCGAG TGCAGAAATA CGGACTCACC TTGCTGTAA 240
 CTACTGATCT TGAGCTCATA AAATACCTAA ATAATGTGGT GGAACAACTG AAAGATTGGT 300
 TATACAGATG TTCAAGTTCAG AAATCTGGTG TAGTTATCTC AAATATTGAA AGTGGTGAGG 360
 TCCTGGAAAG ATGGCAGTTT GATATTGAGT GTGACAGAC TGCAAAAGAT GACAGTGCAC 420
 CCAGAGAAAA GTCTCAGAAA GCTATCCAGG ATGAAATCCG TTCAGTGATC AGACAGATCA 480
 CAGCTACGGT GACATTTCTG CCACGTGTGG AAGTTTCTTG TTCAATTGAT CTGCTGATT 540
 ATACAGACAA AGATTGGTTT GTACCTGAAA AATGGGAAGA GTCGGGACCA CAGTTTATTA 600
 CCAATTCGAA GGAAGTCCGC CTTGTTCTAT TTAATACTAC AATCCACAAA GTAAATAGCA 660
 TGGTGGCCTA CAAAATCTCT GTCAATGACT GAGGATGACA TGAGGAAAAT AATGTAATTG 720
 TAATTTTGA AATGTGTTTT CCTGAAATCA GGTCACTCAT AGTGTATATG TTTTATTTCA 780
 TTGGTTAATT TTTACATGGA GAAACCAAAA ATGATACTTA CTGAAGTGTG TGTAAATGTT 840
 CCTTTATTTT TTTGGTACCT ATTTGACTTA CCATGGAGTT AACATCATGA ATTTATTGCA 900
 CATTGTTCAA AAGGAACCGG GAGGTTTTTT TGTCAACATG GTGATGTATA TTCCTTTGAA 960
 GATAGTAATC GTAGATGGA AACTTGTGTC TATAAGCTA GATGCTTCC TAAATCAGAT 1020
 GTTTTGGTCA AGTAGTTTGA CTCAGTATAG GTAGGGAGAT ATTTAAGTAT AAAATACAAC 1080
 AAGGAAGTCT TAAATATTTA GAATCTTTGT TAAGGTCTCT AAGTAACCT ATAATCTATA 1140
 AACAAATGAA TATGTCTGTA TAGCTCTCTT TGACCTTCAT TTCATGTATA GTTTTCCTTA 1200
 TTGAATCAGT TCCCAATTAT TTGACTTTAA TTTATGTAAC TTGAACCTAT GAAGCAATGG 1260
 ATATTTGTAC TGTTTAATGT TCTGTGATAC AGAACTCTTA AAAATGTTTT TTCATGTGTT 1320
 TTATAAAATC AAGTTTAAAG TGAAAGTGAG GAAATAAAGT TAAGTTTGT TTAATAAAAA 1380
 AAAAAAAAAA

Seq ID NO: 259 Protein sequence:
 Protein Accession #: NP_002349.1

1 11 21 31 41 51
 MALQLSREQ ITLRGSAEIV AEFFSFGINS ILYQRGIYPS ETPTRVQKYG LTLVTTDLLE 60
 LIKYLNNVVE QLKDWLYKCS VQKLVVVISN IESGEVLERW QFIDIECDKA KDDSAPREKS 120
 QKAIQDEIRS VIRQITATVT FLPLLEVSCS FDLIIYTDKD LVVPEKWEES GPQFITNSEE 180
 VRLRSPTITI EKVNSMVAYK IPVND

Seq ID NO: 260 DNA sequence
 Nucleic Acid Accession #: NM_001211
 Coding sequence: 43..3195

1 11 21 31 41 51
 AAAGGCTCTC AGCAGGACGA GGACCTGAGC CAGGAATGCA GGATGGCGGC GGTGAAGAAG 60
 GAAGGGGGTG CTTGAGTGA AGCCATGTCC CTGGAGGGAG ATGAATGGGA ACTGAGTAAA 120
 GAAATGTGAC AACCTTTAAG GCAAGGGCGG ATCATGTCCA CGCTTCAGGG AGCACTGGCA 180
 CAAGAATCTG CTTCTAACAA TACTCTCAG CAGCAGAAAC GGGCATTGTA ATATGAAATT 240

CGATTTTACA CTGGAATGA COCTCTGGAT GTTITGGGATA GGTATATCAG CTGGACAGAG 300
 CAGAACTATC CTCAAGGTGG GAAAGAGAGT AATATGTCAA CGTTATTAGA AAGAGCTGTA 360
 GAAGCACTAC AAGGAGAAA ACGATATTAT AGTGATCCTC GATTTCCTCAA TCTCTGGCTT 420
 AAATTAGGGC GTTATGCAAA TGAGCCTTTG GATATGTACA GTTACTTGCA CAACCAAGGG 480
 ATTGGTGTIT CACTTCTCA GTTCTATATC TCATGGGCAG AAGAATATGA AGCTAGAGAA 540
 AACTTTAGGA AAGCAGATGC GATATTTTCA GAAGGGATTG AACAGAAGGC TGAACCACTA 600
 GAAAGACTAC AGTCCAGCA CCGACAATTC CAAGCTCGAG TGTCTCGGCA AACTCTGTGTG 660
 GCACCTGGA AAGAAGAAGA GGAGGAAGTT TTTGAGTCTT CTGTACCACA ACGAAGCACA 720
 CTAGCTGAAC TAAAGAGCAA AGGGAAGAA ACAGCAAGAG CTCCAATCAT CGGTGTAGGA 780
 GGTGCTCTCA AGGCTCCAAG CCAGAACAGA GGACTCCAAA ATCCATTTC TCAACAGATG 840
 CAAAATAATA TAGAATTAC TGTTTTGTAT GAAAATGCTG ATGAGGCTTC TACAGCAGAG 900
 TTTGCTAAGC CTACAGTCCA GCCATGGATA GCACCCCCCA TGCCCAAGGGC CAAAGAGAAT 960
 GAGCTGCAAG CAGGCCCTTG GAACACAGGC AGGTCCCTGG AACACAGGCC TCGTGCAAT 1020
 ACAGCTTCAC TGATAGCTAT ACCCGCTGTG CTTCCCACTG TCACTCCATA TGTGCAAGAG 1080
 ACTGCAACAC AGCCAGTTAT GACACCATGT AAAATTGAAC CTAGTATAAA CCACATCTTA 1140
 AGCACCAGAA AGCTCGGAAA GGAAAGAAGGA GATCCTCTAC AAAGGGTTCA GAGCCATCAG 1200
 CAAGCGTCTG AGGAGAGAAA AGAGAAGATG ATGTATTGTA AGGAGAAGAT TTATGCAAGGA 1260
 GTAGGGGAG TCTCCTTTGA AGAAATTCGG GCTGAAGTTT TCCGGAAGAA ATTAAAGAG 1320
 CAAGGGAAG CCGAGCTATT GACCACTGCA GAGAAGAGAG CAGAAATGCA GAAACAGATT 1380
 GAAGAGATGG AGAAGAAGCT AAAAGAAATC CAAACTACTC AGCAAGAAAG AACAGGTGAT 1440
 CAGCAAGAAG AGACGATGCC TACAAAGGAG ACAACTAAAC TGCAAATTGC TTCCAGTCT 1500
 CAGAAATAC CAGGAATGAC TCTATCCAGT TCTGTTTGTG AAGTAAACTG TGTGTCAGAG 1560
 GAAACTTCAC TTGCGGAGAA CATTTGCGAG GAACAACCTC ATTCTAAAGG TCCCACTGTA 1620
 CCTTTCTCCA TTTTGTATGA GTTCTTCTT TCAAGAAAGA AGAATAAAAG TCCTCTCTGA 1680
 GATCCCCAC GAGTTTATG TCACCGAAGA CCCTTGCAG TTCTCAAAAC CTCAGAAAGC 1740
 ATCACTCAA ATGAAGATGT GTCTCCAGAT GTTGTGTATG AATTTACAGG AATTGAACCC 1800
 TTGAGCGAGG ATGCCATTAT CACAGGCTTC AGAAATGTAA CAATTTGTCC TAACCCAGAA 1860
 GACACTTGTG ACTTTGCCAG AGCAGCTCGT TTTGTATCCA CTCTTTTCA TGAGATAATG 1920
 TCCTGAAGG ATCTCCCTTC TGATCTCTAG AGACTGTTAC CGGAAGAAGA TCTAGATGTA 1980
 AAGACCTCTG AGGACCGACA GACAGCTTGT GGCATATCT ACAGTCAGAC TCTCAGCATC 2040
 AAGAAGCTGA GCCCAATTAT TGAAGACAGT CGTGAAGCCA CACACTCCTC TGGCTTCTCT 2100
 GGTCTCTCTG CCTCGTGTG AAGCACTCC TCCATCAAA GTCTTCAAA TCCTGAGAAA 2160
 CTAGAACTTA CTAATGAGAC TTCAGAAAC CCACTCAGT CACCATGGTG TTCACAGTAT 2220
 CGCAGACAGC TACTGAAGTC CCTACCAGAG TTAAGTGCCT CTGCAGAGTT GTGTATAGAA 2280
 GACAGACCAA TGCTTAAGTT GGAAATTGAG AAGGAATTG AATTAGGTAA TGAGGATTAC 2340
 TGCAATTAAC GAGAATACCT AATATGTGAA GATTACAGT TATTTCTGGT GGCCTCAAGA 2400
 AACTCTGCAG AATTAACAGT AATAAGGTA TCTTCTCAAC CTGTCCTAGG GGAATTTTAT 2460
 ATCAAGCTCA AGTTAAAGGA ACGTTTAAAT GAAGATTTTG ATCATTTTGG CAGCTGTTAT 2520
 CAATATCAAG ATGCGTGTAT TGTTTGGCAC CAATATATAA ACTGCTTCA CTTCTAGGAT 2580
 CTCTCTCAAC ACAGTGAATA TATTACCCAT GAAATAACAG TGTGATTAT TTATAACCTT 2640
 TTGACAAATG TGGAGATGCT ACACAAAGCA GAAATAGTCC ATGCTGACTT GAGTCAAGG 2700
 TGTCTGATTC TCAGAAACAG AATCCAGAT CCCTATGATT GTAACAGAA CAATCAAGCT 2760
 TTGAAGATAG TGAATTTTC CTACAGTGTG GACCTTAGGG TGCACTGGA TGTTTTACC 2820
 CTCAGCGGCT TTGCGACTGT ACAGATCTCTG GAAGGACAAA AGATCTGGC TAACTGTTCT 2880
 TCTCCCTACC AGGTAGACCT GTTTGGTATA GCAGATTAG CACATTTACT ATTGTTCAAG 2940
 GAACACTTAC AGGTCTTCTG GGATGGGTCC TTCTGGAAC TTAGCCAAA TATTTCTGAG 3000
 CTAAGAGATG GTGAATTTGT GAATAAATTC TTTGTGGGGA TTCTGAATGC CAATGATGAG 3060
 GCCACAGTGT CTGTCTCTGG GGAGCTTGCA GCAGAAATGA ATGGGGTTTT TGACACTACA 3120
 TTCCAAAGTC ACCTGAACAA AGCCTTATGG AAGGTAGGGA AGTTAACTAG TCCTGGGGCT 3180
 TTGCTCTTTC AGTGAAGTAC GCAATCAAGT CTCACAGATT GCTGCCTCAG AGCAATGGTT 3240
 GTATTGTGGA ACATGAAAC TGTATGTGCT GTAATTTAAT TTAGGACACA TTTAGATGCA 3300
 CTACCAATTC TGTCTACTTT TTTGGTACAG GTATATTTTG ACGTCACTGA TATTTTITAT 3360
 ACAGTGATAT ACTTACTCAT GGCCTTGTCT AACTTTTGTG AAGAATCTAT TTATCTTAAA 3420
 CAGACTCATT ACAAATGGTT ACCTTGTAT TTAACCCATT TGTCTCTACT TTTCCCTGTA 3480
 CTTTTCCCAT TGTGAATTTG TAAATGTTT TCTTATGATC ACCATGTATT TTGTAATAA 3540
 TAAATAGTA TCTGTTAAAA AAAAAAAAAA AAAAAAAAAA AAA

Seq ID NO: 261 Protein sequence:
 Protein Accession #: NP_001202

1 11 21 31 41 51
 MAAVKKEGGA LSEAMSELDG EWELSKENVQ PLRQGRIMST LQALAQESA CNNTLQQQKR 60
 AFYEIRFYT GNDPLDVWDR YISWTEQNYF QGGKESNMST LLERAVEALQ GEKRYSDPR 120
 FLNLWLKLR LCNPLDMYS YLHNQIGVVS LAQFYISWAE EYEARENFRK ADAIFQEGIQ 180
 QKAEPLERLQ SQHRQFQARV SRQTLLEALE EEEBEVFESS VPQRSTLAEK KSKGKKKTARA 240
 PIIRVGALK APSONRGLQN PFPQMQNNNS RITVFDENAD EASTAELSKP TVQPWIAAPP 300
 PRAKENELQA GFWNTGRSLR HRPRGNTASL IAVPAVLPSF TPYVEETAQQ PVMTCKIEP 360
 SINHLSTRK PGKEEGDPLQ RVQSHQQASE EKKEKMYCK EKIYAGVGEF SFEEIRAEVF 420
 RKKLKEQREA ELLTSAEKRA EMQKQIEEME KKLKEIQTQ QERTGDQEE TMPTKETTKL 480
 QIASESQKIP TLTSSSVQC VNCCARETSL AENIWQEQPH SKGPSVPSI FDEPLLSEKK 540
 NKSPPADPPR VLAQRRLAV LKTSSESITN EDVSPDVCE FTGIEPLSED AIITGFRNVT 600
 ICPNPEDTCD FARAARFVST PFHEIMSLKD LPSDPERLLP EEDLDVKTSE DQQTACGTIY 660
 SQTLSIKKLS PIIEDSREAT HSSGFGSSSA SVASTSSIKC LQIPEKLELT NETSENPTQS 720
 FWCSQYRRQL LKSLPLSAS AELCIEDRPM PKLEIEKEIE LGNEDYCIKR EYLICEDYKL 780
 FWVAPRNSAE LTVIKVSSQP VPWDFYINLK LKERLNEFD HFCSCYQYQD GCIVWHQYIN 840
 CFTLQDLLOH SEYITHEITV LIYINLLTIV EMLHKAETVH GDLSRCLIL RNRHDPYDC 900
 NKNQALKIV DFSYSVDLRV QLDVFTLSGF RTVQILEGQK ILANCSSPYQ VDLFGIADLA 960
 HLLLFKEHLQ VPWDSFKWL SQNISSELDG ELWNKFPVRI LNADEATVS VLGELAAEMN 1020
 GVFDTTTQSH LNKALWKVGK LTPGALLFQ

Seq ID NO: 262 DNA sequence
 Nucleic Acid Accession #: NM_003784
 Coding sequence: 365..1507

1 11 21 31 41 51

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

GTCTACTTAT CAATAAGCAG CTGCTGTGTC AGAGTGCAGG CTGCACCTTT GGACAGCCTT 60
TAAACTGAA TTCTCAGAAT TTAGAACAA ATTTTGTCT AGAAATGCTG ACTTTGGTTC 120
ATTAGGTAGT GGTAAAAACAG GCTCCCTTCG AAGCTCTCCT TCATCACCTT CCTAAGTGCA 180
TGTCACGGGA AGCTCTCCTT CATCACCTTC CTAAGTGCAT GGGGGAAAAA ACCTAGGGCT 240
CAACAGTCTT GAGAGTGTG GAAACATTTT CTTTGTGAGT GAGAACAGAT CACCTAGAGA 300
AAGGAAACCA GATTCCCATC ACTGCTTCTG GGTATCAGAT GCTAGCGCTG CACTCCATT 360
TGCAATGGCC TCCCTTGCTG CAGCAATGC AGAGTTTTCG TTCAACCTGT TCAGAGAGAT 420
GGATGACAAT CAAGGAAATG GAAATGTGTT CTTTTCCTCT CTGAGCCTCT TCGTGCCCT 480
GGCCCTGGTC CGCTTGGGCG CTCAGATGA CTCCTCTCT CAGATTGATA AGTTGCTTCA 540
TGTTAAACACT GCCTCAGGAT ATGAAATCT TCTAATAGT CAGTCAGGGC TCCAGTCTCA 600
ACTGAAAGA GTTTTTCTG ATATAAATGC ATCCCAAG GATTATGAT TCAGCATTGT 660
GAATGGGCTT TTGTGTAAG AAGTGTATGG CTTTCATAAG GACTACATTG AGTGTGCCGA 720
AAAATTATAC GATGCCAAAG TGGAGCGAGT TGACTTTACG AATCATTAG AAGACACTAG 780
ACGTAATATT AATAAGTGGG TTGAAATGA AACACATGGC AAAATCAAGA ACGTATTGG 840
TGAAGGTGGC ATAAGCTCAT CTGCTGTAAT GGTGCTGGTG AATGCTGTGT ACTTCAAAGG 900
CAAGTGGCAA TCAGCCTTCA CCAAGAGCGA AACATAAAT TGCCATTTC AATCTCCCAA 960
GTGCTCTGGG AAGGCAGTCG CCATGATGCA TCAGGAACGG AAGTTCAATT TGTCTGTAT 1020
TGAGGACCCA TCAATGAAGA TTCTTGAGCT CAGATACAA TGGTGCATAA ACATGTACGT 1080
TCTGCTGGCT GAGATTGAGC TCTCTGAAAT TGAAAAAATA CTGACCTTTC AGAATCTAAT 1140
GGAATGGACC AATCAAGGCG GAATGACCTC TAAGTATGTT GAGGTATTTT TTCCTCAGTT 1200
CAAGATAGAG AAGAATTATG AAATGAAACA ATATTGAGA GGGGTGCTG TGTATATATC 1320
CTTTGATGAA TCCAAAGCAG ATCTCTCTGG GATTGCTTCG GGGGTGCTG TGTATATATC 1320
AAGGATGATG CACAAATCTT ACATAGAGGT CACTGAGGAG GGCACCGAGG CTACTGCTGC 1380
CACAGGAAGT AATATTGTAG AAAAGCAACT CCCTCAGTCC ACGCTGTTTA GAGCTGACCA 1440
CCCATTCTTA TTTGTTATCA GGAAGGATGA CATCATCTTA TTCAGTGGCA AAGTTCTCTG 1500
CCCTTGAAAA TCCAAATGGT TTCTGTTATA GCAGTCCCCA CAACATCAA GRACCACCA 1560
AAGTCAATAG ATYTGRTT AATTGAAAA ATGTGGTGTG TCCTTTGAGT TTATTTCTTC 1620
CTAATATTG TCAAGCAGAT ACATGCTGTA CTTGACCTT CCTAGACACC TGGTTGATTG 1680
TCCTGATCCC TGCTCTTAGC ATTCTACCAC CATGTGCTC ACCCATTTCT AATTTCAATT 1740
TCTTTCTTCC CACGCTCATT TCTATCATT TCCCCATGA CCGCTCTGGA AATTATGGAG 1800
RGTGCTCAAC TGGTAAGGAG AACGTAGAAG TAGCCCTAGG GATCCTTTT GAAACTCTAC 1860
AGTTATCGCA GATATTCTAG CTTCAATTGA AGCAATCTAG GAAATAAGCC CTGCTGCTTT 1920
CTAGAATAA GTGTGAAGGA TAAATTTTCT TTGTTGACCT ATGAAGATT TAGAGTTTAC 1980
CTTCATATG TTGATTTTAA ATCAGTGTAT AATCTAGATG GTAAAAAATG TGAATTTGGG 2040
ATTAGGGACC TACCAAAATA TTTCAATTAAT GCTTCAATT GACAAATTTT GGCCTTTCTT 2100
TGATAAGACA ATATGTACAT GTTTTTCAT ATATTAAGA TCTTTTAAC GTTGGCAGTT 2160
GTTATCTACA GAATCATATT TCATATGCTG TGTAGTTTAT AAGTTTTCCT TCTATTTATC 2220
AGAATAAAGA AATACACAT ACCTGTAAA

Seq ID NO: 263 Protein sequence:
Protein Accession #: NP_003775

45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
MASLAAANAE FCFNLFREMD DNQNGNVFP SLSLPAALA LVRLGAQDD LSQIDKLLHV 60
NTASGYGNSS NSQGLQSQL KRVSFDINAS HKDYDLSIVN GLFAEKVYGF HKDYIECAEK 120
LYDAKVERVD FTHLEDDR NINKWVENET HGKIKNVIGE GGISSSAVMV LVNAVYFKKG 180
WQSAFTSET INCFKSPK SGKAVAMHQ ERKFNLSVIE DPSMKILELR YNGGINMYVL 240
LPENDLSBIE NKLTFNLM EWTNPRMTSK YVEVFFPQFK IEKKNYEMKQY LRLGLKIDIF 300
DESKADLSGI ASGGRLYISR MMHKSIEVET EEGTEATAAT GSNIVEKQLP QSTLFPRADHP 360
FLFVIRKDDI ILFSKVSCEP

Seq ID NO: 264 DNA sequence
Nucleic Acid Accession #: AB052906
Coding sequence: 74-814

60
65
70
75
80
85

1 11 21 31 41 51
AAAACCTTGA GGTGATTCAT CTTCCAGGCT CTCCTTCCAT CAAGTCTCTC CTCCTAGCG 60
CTCTGGGTCC TTAATGGCAG CAGCCGCCGC TACCAAGATC CTTCTGTGCC TCCCGCTTCT 120
GCTCCTGCTG TCCGGCTGGT CCGGGGCTGG GCGAGCCGAC CCTCACTCTC TTTGCTATGA 180
CATCACCGTC ATCCCTAAGT TCAGACCTGG ACCACGGTGG TGTGCGGTTT AAGGCCAGGT 240
GGATGAAAAG ACTTTTCTTC ACTATGACTG TGGCAACAAG ACAGTCACAC CTGTCACTCC 300
CCTGGGGAAG AAACATAAAT TCACAACGGC CTGGAAGCA CAGAACCCAG TACTGAGAGA 360
GGTGGTGGAG ATACTTACAG AGCAACTGCG TGACATTGAG CTGGAGAATT ACACACCCAA 420
GGAACCCCTC ACCCTGCAGG CCAGGATGTC TTGTGAGCAG AAAGCTGAAG GACACAGCAG 480
TGGATCTTGG CAGTTCAAGT TCGATGGGCA GATCTTCTCT CTCTTTGACT CAGAGAAGAG 540
AATGTGGACA ACGGTTTCAT CTGGAGCCAG AAAGATGAAA GAAAAGTGGG AGAATGACAA 600
GGTGTGGGCC ATGTCTCTCC ATTACTTCTC AATGGGAGAC TGTATAGGAT GGCTTGAGGA 660
CTTCTTGATG GGATGAGCA GCACCTTCCA GCCAAGTGCA GGAGCACCAC TCGCATGTC 720
CTCAGGCACA ACCCACTCA GGGCCACAGC CACCACCTCT ATCCTTTGCT GCCTCTCAT 780
CATCTCCCC TGCTTCATCC TCCCTGGCAT CTGAGGAGAG TCCTTTAGAG TGACAGGTTA 840
AAGCTGATAC CAAAAGGCTC GTGTGAGCAC GGTCTTGATC AAACCTGCCC TTCTGTCTGG 900
CCAGCTGCCC ACGACCTACG GTGTATGTCC AGTGGCCTCC AGCAGATCAT GATGACATCA 960
TGGACCAAT AGCTCATTC CTGCTTTGAT TCCTTTTGGC AACAAATTTA CCAGCAGTTA 1020
TACCTAACAT ATTATGCAAT TTTCTCTTGG TGCTACCTGA TGGAAATCCT GCACCTAAAG 1080
TTCTGGCTGA CTAACAAGA TATATCATTT TCTTTCTTCT CTTTGTGTTT GGAAATCAA 1140
GTACTCTTT GAATGATGAT CTCTTCTTGT CAAATGATAT TGTCACTAAA ATAATCACGT 1200
TAGACTTCAG ACCTCTGGGG ATTCTTTCCG TGTCTGAAA GAGAAATTTT AAATATTTA 1260
ATAAGAAAAA ATTTATATTA ATGATTGTTT CTTTGTAGTA TTTATTGTTT TGTACTGATA 1320
TTAAATAAAA GAGTTCTATT TCCCAAAAAA AAAAAAAAAA A

Seq ID NO: 265 Protein sequence:
Protein Accession #: BAB61048.1

1 11 21 31 41 51
 MAAAAATKIL LCLPLLLLLS GWSRAGRADP HSLCYDITVI PKFRPGPRWC AVQGVDEKT 60
 FLHYDCGNTK VTPVSPGLGK LNVTTANKAQ NPVLREVVDI LTELRLDIQL ENYTPKEPLT 120
 LQARMSCBOK AEGHSSGSWQ FSPDQIPLL FDSEKRMWTT VHPGARKMKE KWENDKVVM 180
 SFHYFSMGDC IGWLEDFLMG MDSTLSPSAG APLAMSSGTT QLRATATTLI LCLLLIILPC 240
 FILEPGI

Seq ID NO: 266 DNA sequence
 Nucleic Acid Accession #: XM_084853.1
 Coding sequence: 127-444

1 11 21 31 41 51
 ATTGATGATA TATTAAACGA AATCAAATTT GGTGAATATG TGGACACTGG AAAGCTAATC 60
 GACAAGATCA ACTTACCAGA TTTCTAAAAA GTGTACTCTA ACCACAAGCC ACCTTTTGGT 120
 AACACCATGA GTGGCATCCA CAAGAGCTTT GAGGTGCTCG GTTATACCAA CTCAAAGGG 180
 AAAAAGGCCA TTGGAAGAGA GGACTTCCTG AGACTGCTCG TTACTAAAGG TGAGCATATG 240
 ACGGAGGAGC AGATGTTTGA TTGCTTGTCT TCACCTTTTG GCCTGAATCC CGAGGGATGG 300
 AATTCGAGCG CTGCAACCTG CTCGTCAAA GGTTCAGAAA TTGCTTGA AGAAGAACTT 360
 CCAGACGAAA TCACCTGCAGA AATATTCGCG ACTGAAATTC TTGCTTAAAC CATTTAGAAA 420
 GATTCGGGCC AGGATGGTCA GTGAAGTTAC CAGGAATGTT TAAAGCACA AGGACTTTGG 480
 GTGTGTGTGC ATGCACATGT GTGTGTTTC CATGAGGCAC TGCTTTTAT GCATTTCCCT 540
 CCCCCCTCTC ATCTTTAGAA CATTTAGACA TTAAAGCAAG TTTCTGGTGA GCAATG

Seq ID NO: 267 Protein sequence:
 Protein Accession #: XP_084853.1

1 11 21 31 41 51
 MSGIHKSFEV LGYTNKSGKK AIRREDFLRL LVTKGEHMT EEMLDCEASL PGLNPEGWKS 60
 EPATCSVKGS BICLEELPD EITAEIFATE ILGLTISEDS GQDQG

Seq ID NO: 268 DNA sequence
 Nucleic Acid Accession #: NM_001898
 Coding sequence: 57-482

1 11 21 31 41 51
 GGCTCTCACC CTCTCTCTCT GCAGCTCCAG CTTTGTGCTC TGCTCTGAG GAGACCATGG 60
 CCCAGTATCT GAGTACCCTG CTGCTCCTGC TGGCCACCCT AGCTGTGCGC CTGGCCTGGA 120
 GCCCAAGAGA GGAGGATAGG ATAATCCCGG GTGGCATCTA TAACGCAGAC CTCAATGATG 180
 AGTGGGTACA GCGTGCCCTT CACTTCGCCA TCAGCGAGTA TAACAGGCC ACCAAGATG 240
 ACTACTACAG AGTTCGCTG CGGTACTAA GAGCCAGGCA ACAGACCGTT GGGGGGGTGA 300
 ATTACTTCTT CGACGTAGAG GTGGGCGCA CCATATGTAC CAAGTCCCAG CCCAACTTGG 360
 ACACCTGTGC CTTCATGAA CAGCCAGAAC TGCAGAAGAA ACAGTTGTGC TCTTTCGAGA 420
 TCTACGAAGT TCCTTGGGAG AACAGAAGGT CCCTGGTGAA ATCCAGGTGT CAAGAATCCT 480
 AGGGATCTGT GCGAGGCCAT TCGCACCAGC CACCACCCAC TCCACCCCC TGTAGTGCTC 540
 CCACCCCTGG ACTGGTGGCC CCACCCCTGC GGGAGGCCTC CCATGTGCTC TGCCTCAAGA 600
 GACAGACAGA GAAGGCTGCA GGAGTCTTT GTTGCTCAGC AGGGCGCTCT GCCCTCCCTC 660
 CTCTCTTCTT GTTCTTAATA GCCCTGGTAT ATGGTACACA CCCCCCACC TCCTGCAATT 720
 AACAGTAGC ATCGCC

Seq ID NO: 269 Protein sequence:
 Protein Accession #: NP_001889.1

1 11 21 31 41 51
 MAQYLSTLLL LLATLAVALA WSPKEEDRII PGGIYNADLN DEWQORALHF AISEYNKATK 60
 DDYRRPLRV LRARQQTIVG VNYFFDVEVG RTICTKSQPN LDTCAFHEQP ELQKKQLCSF 120
 EIYEVWENR RSLVKSRCQS S

Seq ID NO: 270 DNA sequence
 Nucleic Acid Accession #: XM_093210
 Coding sequence: 13-1854

1 11 21 31 41 51
 ATGGCAAGCG CCGGAATCTC CTCAGCTGCC GTTTCACAAA AGAGGTACCA GGTCCGCACC 60
 AAACGAGCAC ACAAGCAGCA CCAGGAGCTG CAGAAGAAGG AGGCGGCAGC GATGGACCAG 120
 GGCAGAGGGA ATGGGGGAGG GGCATCCTAC CCCATATCTG AGGTGCGACT GCGGGACGTA 180
 GAGCGGACTG GGCCTTTCCC GTTGGCGCGT GGCCTCAATC AGGACTTCTT GCCACGTGTC 240
 GCCTTCAAAA CGGTAAGAGC TGCAACTGAA CGTGTGAGAC ATGGTGCAGA TAGGCTGAGA 300
 GGCAGCGGGA GAGATGCCCA TGAACCTAAG TACCGGAGCA CGCCTCCAC TTCTACCAAC 360
 ACGAGTAACA CCGCCCCCAG GGGACCGCTC TCGAGGTCCC CCAAGCCAAG GACGCAAGGA 420
 GGAACGCCCC GGCAGCGGCG CAGCAGCGGC GGGCACCAGC CCAATGGCCA CGGAATCTAG 480
 CACTGGCAGT CGGCCCTCTC CACACCGCAG GCGTGCAAGT TGGCCGACGG AGCCTCCCGG 540
 GCCGAGSACC CAGCTAGGCC GTCACCCCGG TTGCTCCAC GGGGAAGGGC ACCAGGCAAA 600
 CTGCCCAAGG CCCGAGCCCC AGGCTCCCTG GCGGAGGCCT CCGCTGGTCC CGCCAGATC 660
 ATGGCCGCA CCAGGCTCCC GAGCCATGGC TTCTGTGCTG GGAACGCCCC GGCCTCCTGG 720
 CTGTCCAGCT AG

Seq ID NO: 271 Protein sequence:
 Protein Accession #: XP_093210

1 11 21 31 41 51

MLRHGEQKRR RARKNDLFLP TCAFKTVRAA TERVRHGADR LRGGGRDAHE LKYPDTPSTS 60
TTTSNTAPYG PLSRSPKPRP QGGTPRRRPA AAGTRANGHG TQHWQSALLT PQACSVADGA 120
5 SRADPARPS PRLPREGAP GKLPKAPSPG SLAEASAGLL AHVRLQNADA QRVSIQALP 180
PNSVGRKEE RFGAGQQRRA PAPMATELST GSRPSSHRRR AWPTEPPGP RTQLEPSPRL 240
LPREGAPGKL PKAPSPGSLA EASAGPAQIM AATRLPSRGP LSGNGPASWL SS

Seq ID NO: 272 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..732

1 11 21 31 41 51
GGATACGTG TCACTCAAAAG TAATGGGAGG GAGAGAGAAC AGGGAGGGTA GGGATGCTTT 60
15 TGAAAAAGCT TTTTTCCTCA CTTTAACTT GCTTTAGCGT TAAGAGTACT TACCAGCTAA 120
TAATGTGGAG GAAATATTCT TTTCTCATG GAGATTACAG AATATATCTA TTCTCTTGA 180
ATACCCACTT GAAGCCTCTG TAGAAATGTC TCGTCTCCCG GTTGTATTTC TAAACCTAC 240
ATGATTTTGT CTTGTTTCTG CAGTGAGAAA TTACATCCAT AGCAAAGACA AAAGTCTTTT 300
TAAATTATT TTATTATCTT TTCATATAGT TCTTACAATT TCTAAAAAAT TAACACTCAT 360
20 TTAGTATCAC AATTATATGG AGAGGGTTT TTGTATTTT AAGCATATGT GGCTTATATA 420
AAATTGCAG AAGTCATAGG ACTGTCATGT ATTGCAGCTC TGAGAACCAA TGCCGTGAAC 480
TTAAGCC

Seq ID NO: 273 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
MGGRENREGR DAFKAPFFPT FNLL

Seq ID NO: 274 DNA sequence
Nucleic Acid Accession #: NM_003976.2
Coding sequence: 299-961

1 11 21 31 41 51
CTCTGAGCTT CTCTGAGCCT TGTGTGCTCA TCTGGAAAAA GGGGATTAAA CCAITTAACCT 60
40 CATGGAGTTG TGAAAGAATA GCTGCAAAAG ACCTAACACA TAGTAAGGTT CCCAGTGCAG 120
CTACTCTGCT TGGGTTGAGT CTAGCTGTGT AGGCCCTCTG TTCTCACCTT GGAGAAACTG 180
GGGTGGCAGG CCGGTCCTCC ACAAAAGATA ACTCATCTCT TAATTGCAA GCTGCCTCAA 240
CAGGAGGGTG GGGGAACAGC TCAACAATGG CTGATGGGCG CTCTGGTGTG TGATAGAGAT 300
GGAACCTGGA CTTGGAGGCC TCTCCACGCT GTCCCACTGC CCTGGGCTA GCGCGCAGC 360
TGCCCTGTGG CCCACCTGCG CCGCTCTGGC TCTGCTGAGC AGCGTCGAG AGGCCTCCCT 420
45 GGGCTCCCGG CCGCGCAGCC CTGCCCCCGG CGAAGGCCCC CCGCTGTGCC TGGGCTCCCT 480
CGCGCGCCAC CTGCGGGGGG GAGCGACGCG CCGCTGGTGC AGTGAAGAG CCGCGCGGCC 540
GCGCGCGCAG CCTTCTCGCG CCGCGCCCCG GCGCCTGCA CCCCCTCTG CTCTCCCGG 600
CGGGGGCGCG GCGCGCGCGG CTGGGGGGCC GGGCAGCGCG GCTCGGCAG GGGGGGCGG 660
GGGTGCGCGC CTGCGCTCGC AGCTGGTGCC GGTGCGCGCG CTCGGCCTGG GCCACCGCTC 720
50 CGACGAGCTG GTGCGTTTCC GCTTCTGCG CCGCTCCTGC CGCGCGCGCG GCTCTCCACA 780
CGACCTCAGC CTGGCCAGCG TACTGGGCGC CGGGGCCCTG CGACCGCCCC GGGGCTCCCG 840
GCCCGTCAGC CAGCCTCTGC GCGCAACCC GCGCTACGAA GCGGTCTCTT TCATGGAAGT 900
CAACAGCACC TGGAGAACC TGGACCGCCT CTCGCCACCC GCCTGCGGCT GCCTGGGCTG 960
AGGGCTCGCT CCAGGCTTTT GCAGACTGGA CCTTACCGG TGGCTCTTCC TGCTTGGGAC 1020
55 CTTCCCGCAG AGTCCCACTA GCCAGCGGCC TCAGCCAGGG ACGAAGGCTT CAAAGCTGAG 1080
AGGCCCTAC CCGTGGGTGA TGGATATCAT CCCCAGAAC GTGAGGGGAC AACTGACTAG 1140
CAGCCCCAG GCGCTCCAG TCGGATCCC AGCCTAAAAG ACACAGAGA CCTCAGCTAT 1200
GGAGCCCTTC GGACCCACT CTCACAGACT CTGGCACTGG CCAGGCTCG AACCTGGGAC 1260
CCCTCTCTG ATGAACATA CAGTGGCTGA GGATCAGCC CCCGCCAGG CCTGTAGGG 1320
60 ACAGCATTG AAGGACATAT ATTGCAATTG CTTGGTTGAA AGTGCGTGTG CTGGAACCTG 1380
CCTGTACTCA CTATGGGAG CTGGCCCC

Seq ID NO: 275 Protein sequence:
Protein Accession #: NP_003967.1

1 11 21 31 41 51
MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60
70 PAGHLPGGRT ARWCGRARR PPPQSRPAP PPPAPPSALP RGGRAARAGG PGSRARAAGA 120
RGCRLRSQLV PVRLGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLIG AGALRPPPGS 180
RPVSQPCRRP TRYAVSPMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 276 DNA sequence
Nucleic Acid Accession #: NM_057091.1
Coding sequence: 783-1445

1 11 21 31 41 51
ACTGGCCGCT GAGAGAAGAA TCGGGTGGAG CAGAGAGCAG CTGCTGCAGG GCAGACAGCC 60
80 GGACCCCAA ATCTGCAGT ACCAGCAGTC AGCCGCCCA CAGAGGAGCC GGCTTACCCC 120
TGGCTCCCGC CCTCACTCA CTTTCTCCCG CCTCGGCCCG GGCCTCCCA CTCTCTACTT 180
CGCGTGTCTA CAACTCAAC TCCCGTTTC CGTGCTCTC CACGCTCGA GTTCTCTACT 240
CTCCATATCC GAGGGGCCCC TCCAGCATC TACCCCTCT CCAACTCTG GGGACCTAGC 300
CAAGCTAGGG GGGACTGGAT CCGACGGGTG GAGCAGCCAG GTGAGCCCGC AAAGGTGGGG 360
85 CGGGCAGGG GCGCTCCAG CCCCACCCCG GGATCTGGTG ACGCTGGGGC TGGAAATTGA 420
CACCGAGGG CTGCGCGCGC GGGCAGGAG CTGCTGAGGG ATGGAGTTGG GCCCGGCCCG 480
CAGACAAGGC CCGGGGGCTC CGCCAGCAGC AGGTCCCTCG GGGCCAGCC CTGCTGCCA 540

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60

CCGGGGCTG GAGCCACAC CCGAGGGTG CAGACTGGCT GCCAAGGCCA CACTTTTGGC 600
 TAAAAGAGGC ACTGCCAGGT GTACAGTCCT GGGCATGGCG TGTTTGAGCT TCGGGGGAGA 660
 GCCCAGCACT GGTCCCGGGA AAGGTGCCTA GAAGAACAAAG GTGCAGGACC CCGTGCTGCC 720
 TCAACAGGAG GGTGGGGGAA CAGCTCAACA ATGGCTGATG GGGCTCTCTG GTGTGTATAG 780
 AGATGGAAT TGGACTTGA GGCCTCTCCA CGCTGTCCCA CTGCCCTGG CTTAGGGGGC 840
 AGCCTGCCCT GTGGCCACC CTGGCCGCTC TGGCTCTGCT GAGCAGGCTC GCAGAGGCTC 900
 CCCTGGGCTC CGCGCCCGC AGCCCTGCCC CCGCGAAGG CCCCCCGCT GTCTTGGCT 960
 CCGCGCCCG CACCTGCGG GGGGAGCGCA CGGCCGCTG GTGCAGTGA AGAGCCCGGC 1020
 GCGCGCCCG CGACCTTCT CGGCCCGCG CCGCGCGCTG TGCACCCCA TCTGCTCTTC 1080
 CCGCGGGGG CCGCGCGCG CGGCTGGGG GCGCGCTCG GCAGCGGGG 1140
 CGCGGGGCT CGCGCTGCG TCGCAGCTGG TGCCGCTGG CGCGCTCGG CTGGGCCACC 1200
 GCTCCGACGA GCTGGTGGT TTCCGCTTCT GCAGCGGCTC CTGCGCGCG GCGCGCTCTC 1260
 CACACGACCT CAGCCTGGCC AGCCTACTGG GCGCGGGGC CCTGCGACG CCCCCGGCT 1320
 CCGCGCCCGT CAGCCAGCC TGCTGCCGAC CCAACGCTA CGAAGCGGT TCCTTCATGG 1380
 ACGTCAACAG CACTTGGAGA ACCGTGGACC GCCTCTCGC CACGCGCTG GGTGCGCTG 1440
 GCTGAGGGCT CGCTCAGGG CTTTGAGAC TGGACCTTA CCGTGTGCTC TTCTGCTCTC 1500
 GGACCTCTCC GCAGAGTCCC ACTAGCCAGC GGCCTCAGC AGGACGAAG GCCTCAAAGC 1560
 TGAGAGGCC CTACCGGTGG GTGATGGATA TCATCCCGA ACAGTGAAG GSACAACCTG 1620
 CTAGCAGCCC CAGAGCCCTC ACCCTGCGGA TCCAGCCTA AAAGACACCA GAGACCTCAG 1680
 CTATGGAGCC CTTGGAGACC ACTTCTCACA GACTCTGCGA CTGCCAGGC TCGAACCTG 1740
 GGACCCCTCC TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCCCGCC CAGGCCCTGT 1800
 AGGACAGCA TTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAGTGCC TGTGCTGGAA 1860
 CTGGCTGTA CTCACCTATG GAGCTGGCC CC

Seq ID NO: 277 Protein sequence:
Protein Accession #: NP_003967.1

1 11 21 31 41 51
 MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60
 PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPSALP RGGRAARAGG PGRSARAAGA 120
 RGCRRLRQLV PVRALGLGHR SDELVRFRFC SGSCRRRASP HDLSLASLLG AGALRPPPGS 180
 RPVSQPCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 278 DNA sequence
Nucleic Acid Accession #: NM_057160.1
Coding sequence: 1-714

1 11 21 31 41 51
 40 ATGCCCGGCC TGATCTCAGC CCGAGGACAG CCGCTCCTTG AGGTCTTCC TCCCAAGCC 60
 CACCTGGGTG CCCTCTTCT CCCTGAGGCT CCACTTGGTC TCTCCGCGCA GCCTGCCCTG 120
 TGGCCACACC TGGCCGCTCT GGCTCTGCTG AGCAGGCTCG CAGAGGCTC CCTGGGCTCC 180
 GCGCCCGCA GCCCTGCCCC CCGGGAAGGC CCGCGGCTG TCTGGGCTC CCGCGCGGC 240
 CACCTGCCGG GGGGAGCAC GGCCTGCTGG TGCACTGGA GAGCCCGCG GCGCGCGCG 300
 45 CAGCTTCTC GCGCGCGCC CCGCGGCTC GCACCCCAT CTGCTCTTC CCGCGGGGG 360
 CCGCGCGCG GGGCTGGGG CCGGGGAGC CGCGCTCGG CAGCGGGGG GCGGGGCTGC 420
 CGCTTGGCT CCGAGCTGGT GCGGTGCGC GCGCTGGCC TGGGCCACG CTCCGACGAG 480
 CTGGTGGCT TCCGCTTCT CAGCGGCTCC TGCGCGCGC CGCGCTCTC ACACGACCTC 540
 AGCCTGCCCA GCTTACTGGG CGCGGGGCC CTGCGACCG CCGCGGCTC CCGCGCGCTC 600
 50 AGCCAGCCCT GCTGCCGACC CACGCGCTAC GAAGCGGTCT CCTTCATGA CGTCAACAGC 660
 ACCTGGAGAA CGTGAGACC CCTCTCCGCC ACCGCTGCG GGTGGCTCT TCTGCTGGG CTGAGGGCTC 720
 GCTCCAGGGC TTGTCAGACT GGACCTTAC CGTGGCTCT TCTGCTGGG GACCTTCCG 780
 CAGAGTCCCA CTAGCCAGG GCCTCAGCCA GGGACGAAG CCTCAAAGT GAGAGGCCCC 840
 TACCGTGGG TGATGGATAT CATCCCGAA CAGGTGAAG GACAACTGAC TAGCAGCCCC 900
 55 AGAGCCCTCA CCGTGGGAT CCCAGCCTAA AAGACACAG AGACCTCAGC TATGGAGCCC 960
 TTGGGACCCA CTTCTCAGC ACTCTGGCAC TGGCCAGGCC TCGAACCTGG GACCCCTCT 1020
 CTGATGAACA CTACAGTGGC TGAGGCATCA GCGCCCGCC AGGCCTGTA GGGACAGCAT 1080
 TTGAAGGACA CATATTGCA TTGCTTGGTT GAAAGTGCT GTGCTGGAAC TGGCCTGTAC 1140
 TCACTCATGG GAGCTGGCCC C

Seq ID NO: 279 Protein sequence:
Protein Accession #: NP_476501.1

1 11 21 31 41 51
 65 MPGLISARQ PLLEVLPPQA HLGALFLPEA PLGLSAQPAL WPTLAALALL SSVAEASLGS 60
 APRSPAPREG PPPVLSPAG HLPGGRTARW CSGRARRPPP QPSRPAPPPP APPSALPRGG 120
 RAARAGGPGS RARAAGARG RLRSQLVPR ALGLGHRSD LVRFRFCSGS CRRARSPHDL 180
 SLASLLGAGA LRPPGSRPV SQPCCRPTRY EAVSFMDVNS TWRTVDRLSA TACGCLG

Seq ID NO: 280 DNA sequence
Nucleic Acid Accession #: NM_057090.1
Coding sequence: 29-715

1 11 21 31 41 51
 75 CTGATGGGCG CTCCCTGGTG TGATAGAGAT GGAACCTGGA CTTGGAGGCC TCTCCACGCT 60
 GTCCACTGC CCTTGGCTA GCGGCGAGGC TCCACTTGGT CTCTCCGCG AGCCTGCCCT 120
 GTGGCCACAC CTGGCGGCTC TGGCTCTGCT GAGCAGGCTC GCAGAGGCTC CCTTGGGCTC 180
 80 CCGCGCCCGC AGCCTTGGC CCGCGGAAG GCGCGGCTG GTGCACTGGA AGAGCCCGC GCGCGCGCC 240
 CCACTTGGC GGGGAGCGCA CCGCGGCTG GTGCACTGGA AGAGCCCGC GCGCGCGCC 300
 GCAGCTTCT CCGCGCGCG CCGCGCGCC TGCACCCCA TCTGCTCTC CCGCGGGGG 360
 CCGCGCGCG CCGCTGGGG GCGCGGCGC CCGCGCTCG GCAGCGGGG CCGCGGGCTG 420
 CCGCTGCGC TCGCAGCTGG TGCGGTGGC CCGCTCGGC CTGGCCACC GCTTCCAGCA 480
 85 GCTGGTGGT TTCCCTTCT GCAGCGGCTC CTGCGCGCG GCGCGCTCT CACACGACCT 540
 CAGCTTGGC AGCCTACTGG GCGCGGGGC CCGCGCGGC CCGCGCGCT CCGCGCGCT 600
 CAGCCAGCCC TGCTGCCGAC CCAACGCTA CGAAGCGGT TCTTCATGG ACCTCACAG 660

WO 02/086443

PCT/US02/12476

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

CACCTGGAGA ACCGTGGACC GCGTCTCCGC CACGCGCTGC GGCTGCTGG GCTGAGGGCT 720
CGCTCCAGGG CTCTTGACAC TGSACCCCTTA CCGGTGGCTC TTCTGCTCTG GGACCTCCC 780
GCAGAGTCCC ACTAGCCAGC GCGCTCAGCC AGGGACGAAG GCCTCAAAGC TGAGAGGCCC 840
CTACCGGTGG GTGATGGATA TCATCCCGGA ACAGGTGAAG GGACAACTGA CTAGCAGCCC 900
CAGAGCCCTC ACCCTGCGGA TCCCGCCTA AAAGACACCA GAGACCTCAG CTATGGAGCC 960
CTTCGGACCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTGAAACCTG GGACCCCTCC 1020
TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCCCGCC CAGGCCCTGT AGGGACAGCA 1080
TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAAGTGCC TGTGCTGGAA CTGGCTCTGA 1140
CTCACTCATG GGAGCTGGCC CC

Seq ID NO: 281 Protein sequence:
Protein Accession #: NP_476431.1

1 11 21 31 41 51
| | | | |
MELGLGLST LSHCPWPRRQ AFLGLSAQPA LWPTLAALAL LSSVAEASLG SAPRSPAPRE 60
GPPPLVASPA GHLPGGRTAR WSGRRARRPP PQPSRPAPPP PAPPSPALPRG GRAARAGGGP. 120
SRARAAGARG CLRLRSQLVFV RALGLGHRSD ELVRFRCPSG SCRRARSPHD LSLASLLGAG 180
ALRPPPGSRP VSQPCCRPFR YEAVSFMDVN STWRTVDRLS ATACGCLG

Seq ID NO: 282 DNA sequence
Nucleic Acid Accession #: Eos sequence

1 11 21 31 41 51
| | | | |
CTACTGCACC TGCCTCTGT TTCCTTTGGA AATCTCTTAC CTTTCATTAG GGTTCCTTTC 60
ATAGCAATTT CCTTTGGTTT TTAAGACTTC TACATTGCTT TTTCTTTTAT TATCTGTGCT 120
CCGTGAACCT TATGAATGCT GCTTAAAAAT AATGTCAAAA TATGTTTATG CTGCTTACTC 180
AGGTAACGTT TTTCTTTGCT CTATCTTGG TTTCCATATA CTATTTTGGG TTTTCTGTGA 240
GATCTAATCA ATGATCTAGT CAGAAGCTAC TTCCTGGCT AACAGTGATC ATGTTCTATG 300
GCTAAAAATG AACTTGAAC ACAGAAAGTAG TGGTTGGTCC AGTTTGAAAG CTCTTATTAG 360
TATTTCTTAT CTTGGCTGTA ATAATAGCCA TTATTTGTTA TGCCTTTGTT ATGTAGCAGA 420
CACTCTTAAG GATTTTATGT GTATTATTCA AATGCTATT ACTGTTCTTT TTATAGTTGA 480
GAATCTCAGG ATACCTACAT TTATCACTTT TTCAATATAT ATGTATTCTT TATT

Seq ID NO: 283 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 564-1481

1 11 21 31 41 51
| | | | |
GAGACTTTTA ATCATCTATC CCTTGTGCTT TACGACAGCC CTACAATACA CTAGAGGCTT 60
CAAGAGGCTC AAAAATTCAC ATGTGTAGAC AAATTAGGTC CCTTAAGATG CCAGGCAAAAC 120
GAAGTGCTAC CAACACACGC AATGACTGTC CTAAGATGTC GTTCTGGGAT ACACCTGTAA 180
ACTTGATGCA AGTTCCCTCC CCTCTCTCA AAATATATCG ACTTGTGCTG AAAGAAATCA 240
CGACOGATGC TCACAATCTC GACCTCGTAA TTATATAGGG GGTGGTTTGT GTTCTGCGT 300
CTTTCCCTGA TTCACTGGCA GGTAAACATAT TTCATGTACA AAATGAACCTG CAACACCACG 360
GCAACAAGAG GACAGGCCCT CAAAGTTGTC GGTAGGGAGC CAGGACCCCG CCACTGGCGT 420
GGGGAGACAC CGTACTAAAC AAGCTTGCAA ACAGCAGGCA CCTTCTGCCC ACTGAGGAGG 480
AAGGGCTGCG TAAGGGAGGC CGGGGCGGAG GAAGCCAAGC TGTGACGGCC CTGACAAAGT 540
CCTCCCGGCC TCCACCGCTC GCCATGGCAA CGCGGGTCT GTGCTGGCCG GGATTGGCCG 600
GCTTGGCGCG CGCAGGGCCC GCTGGGAAAG CGCGTCCC CGCGGCTCC GCCAGTTTGA 660
ACTTGGCGGG CCAGATGTGG GCGGCGGGGC GCTGGGGGCC TACTTTTCCC TCTTCTACG 720
CCGTTTCTC TGCTGACTGC AGACCCAGGT CTCGGCCCTC CTGGAAGTCC TGCTCAGTCC 780
CTATGACGCG CGCACGTGGG CAGGGGCTGG AGGTGGTGCG CTCGCGCTCG CCGCCGCTGC 840
CGCTGAGCTG CAGCAATTC ACCAGGTGCG TGTGTCTCC CCTTGGCCAC CAGAGCTTCC 900
AGTTTGACGA GGACGACGGT GACGGGGAGG ATGAGGAAGA CGTGGATGAT GAGGAAGACG 960
TGATGAAGA TGCCCATGAT TCAGAGGCCA AAGTGGCGAG CCTGAGAGGA ATGAGTTTAC 1020
AGGGTGGCG CAGCACTCAG GTTGAATCAG AAAATAACCA AGAAGAACAG AAACAGGTGC 1080
GCTTACCAGA AGCGCCCTG ACACCATGGG AGGTGTGGTT TATTGGCAAA GAAAAAGAAG 1140
AACGTGACCG GCTGCAACTG AAAGCTCTAG AGGAATTAAT TCACAACTA GAAAAAGAA 1200
AAGAAATGGA AGAAGCTGAA AAAAGAAAGA TAATTGCTGA AGAAAGCAC AAGGAATGGG 1260
TTCAGAAAAA GAATGAGCAA AAAAGAAAAG AAAGAGAAC AAAAAATTAAT AAAGAAATGG 1320
AGGAAAAAGC AGCAAGGAA CTGGAGAAAG AATACTTGCA AGAAAAAGCA AAAGAAAAAT 1380
ATCAAGAATG GTTAAAGAAA AAAAATGCTG AAGAATGTGA GAGGAAGAAG AAAGAAAAAG 1440
AAAAACAACG CAAGCTGAAA TACAGGAGAA AAAGGAAATA GCAGAAAAAA AGTTTCAAGA 1500
ATGGTTGGAA AATGCGAAAC ATAAACCTCG TCCAGCTGCA AAGAGCTATG GTTATGCCAA 1560
TGGAAAACTT ACAGGTTTTT ACAGTGGAAA TTCTATCCA GAACCAAGCT TTTATAATCC 1620
AATTCCGTGG AAACCAATTC ATATGCCACC TCCCAAGAA GCTAAGGATC TATCAGGAAG 1680
GAAGAGTAAA AGACCTGTGA TAAGTCAGCC ACACAAGTCA TCATCTCTGG TAATTCATAA 1740
AGCCAGGAGC AATCTTTGCC TTGGAACTCT GTGCAGAAAT CAAAGATAGC GTATGTGGAA 1800
AATAACATGC TTTTATCTGG AGCTATTTAA TTTAAAAATC AGAAATGTT TTTTACTGCT 1860
CAGTCAATAA CTCACACTT AATGTGATTA TTGACAAATA GCAATTTTGT CATTGTGATA 1920
TGGATCTCTT AGAGTTGAGG AAGATATTTT CTGGATTTTG GTTTTATAAA ACTTTTAAAG 1980
GTTGATCTTG GCATGTTGTT TTGCAGAAAT AGTGGCTGAA TATGTAAGAA TTGTGTTTGT 2040
ATTATGCTTG TATTAAGAGT AACTGTAAAT ACCAATAAAA CTAACAATTT TTCTTG

Seq ID NO: 284 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
| | | | |
MATRGLCWPG LAGLARAGPA GKARPRRGS SLNLAGQMHA AGRWGPTFPS SYAGFSADCR 60
PRSRPSSDC SVPMTGARGQ GLEVVRSPSP PLPLSCSNST RSLSPPLGHQ SFQFDEDDGD 120
GEDEEDVDDE EDVEDAHDS EAKVASLRGM ELQGCSTQV ESENNEEQK QVRLPESRLT 180
PWEVNFQKE KEERDRQLK ALEELNQLE KRKEMESREK RKI IAEKHK EWWQKNEQK 240
RKEREQKINK EMEEKAARKEL EKEYLQEKAK EKYQSWLKKK NAECEKRRKK EKKNNSKLRY 300

RRKKR

Seq ID NO: 285 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1-1746

1. 11 21 31 41 51
10 ATGCCACTGA AGCATTATCT CTTTGTGCTG GTGGGCTGCC AAGCCTGGGG TGCAGGGTTG: 60
GCCTACCATG GCTGCCCTAG CGAGTGTAAC TGCTCCAGGG CCTCCAGGT: GAGTGCACCC: 120
GGGGCAGCA TTGTGGCGGT GCCCACCCCT CTGCCCTGGA ACGCCATGAG CCTGCAGATC 180
CTCAACACGC ACATCACTGA ACTCAATGAG TCCCGTTC TCATATCTC AGCCCTCATC 240
GCCCTGAGGA TTGAGAAGAA TGAGCTGTGC CGCATCACGC CTGGGGCCTT: CGAAACCTG 300
GGCTGCTGC GCTATCTCAG CCTGCCAAC AACAGCTGC AGGTTCTGCC CATCGGCTC 360
15 TTCCAGGGCC TGGAGAGCCT TGAGTCTCTC CTTCTGTCCA GTAACCACT: GTTCAGATC 420
CAGCGGCGCC ACTTCTCCCA GTGCAGCAAC CTCAGGAGC: TGCAGTTGCA CGGCAACCAC 480
CTGGAATACA TCCCTGACGG AGCCTTCGAC CACCTGGTAG GACTCAGAA: GCTCAATCTG 540
GGCAGAATA GCCTCACCCA CATCTCACCC AGGGTCTTCC AGCAGCTGGG: CAATCTCCAG: 600
GTCTCTCGGC TGTATGAGAA CAGGCTCACG GATATCCCA: TGGGCACTT: TGTAGGGCTT: 660
20 GTTAACCTGC AGGAACCTGGC TCTACAGCAG: AACAGATTG GACTGCTCTC: CCTGCTCTC 720
TTCCACAACA ACCACAACCT: CCAGAGACTC TACCTGTCCA: ACAACCATC: CTCACAGCTG 780
CCACCCAGCA TCTCTATGCA GCTGCCCCAG: CTCACCGTTC: TTAATCTCTT: TGGGAATTC: 840
CTGAAGGAG TCTCTCTGGG GATCTTCGGG CCAATGCCCA ACCTGCGGGA: GCTTTGGCTC 900
25 TATGACAACC ACATCTCTTC TCTACCCGAC AATGTCTTCA GCAACCTCCG: CCAATGTCAG 960
GTCTCTGATC TTAGCCGCAA TCAGATCAGC: TCTATCTCCC CGGGTGCCTT: CAACGGGCTA 1020
ACGGAGCTTC TGGAGCTGTC CCTCCACACC AACGCACTGC AGGACCTGGA: CGGGAATGTC 1080
TTCCGATGT TGGCAACCTC GCAGAACATC: TCCCTGCAGA ACAATCGCCT: CAGCAGCTC 1140
CCAGGGAATA TCTTCGCAA CGTCAATGGC CTCATGGCCA: TCCAGCTGCA GAACAACCA: 1200
CTGGAGAATC TGCCCTCGG: CATCTTCGAT CACCTGGGGA AACTGTGTGA: GCTCGGGCTG: 1260
30 TATGACAATC CCTGAGGTC: TGACTCAGAC ATCCTTCGCG TCCGCACTG: GCTCCTGCTC 1320
AACAGCCTTA GGTAGGAGC: GGCACACTGA CCTGTGTGTT: TCAGCCAGC: CAATGTCCGA: 1380
GGCCAGTCCC TCATATCAT CAATGTCAAC GTTGCTGTTC: CAAGCGTCCA: TGTCCCTGAG 1440
GTGCTAGTT ACCCAAGAAC ACCATGGTAC CCAGACACAC CCAATACCC: TGACACCACA 1500
35 TCCGTCTCTT CTACCACTGA: GCTAACCCAGC CCTGTGGAAG: ACTACACTGA: TCTGACTACC 1560
ATTGAGTCA: CTGATGACCG: CAGCGTTTGG: GGCATGACCC: AGGCCAGAG: CGGCTGGCC 1620
ATTGCGGCA TTGTAATTTG CATTGTGCGC CTGGCTGCTC: CCTGGCTGCTC: CTGGCTGCTC 1680
TGTGCTGCT GCAAGAGAG: GAGCAAGCT: GTCCGTGATG AGATGAAGGC: ACCCAATGAG 1740
40 TGTAAAGAG GCAGGCTGGA GCAGGCTGG: GGAATGATGG GACTGGAGGA: CTTGGGAATT 1800
TCATCTTTCT GCCTCCACCC CTGGGTCCAT: GGAGCTTTCC CGTGATTGCT: CTTTCTGGCC: 1860
CTAGATAAG: GTGTGCTTAC: CTCTTCTTGA: CTTGCTGAT TCTCCGTCAG: AGAAGCAGGT: 1920
CGTGCGGAC CTTCTACAA TCAGGAAGAT: AGATCCAAT GGCATAGGA: AAAGCCCTGG: 1980
GGATTTCCGA: TCTATACCCC TGGGCTTCT: TCGAGAGGCG: TCTTCTTCCA: AATCTCCCC 2040
45 ACCTGTCTC CAAGAACAGC CTTCCCTGCG CCGAGGCCCC CTCCGGCCTC: CTGTAGACTC: 2100
AGTTAGTCCA CAGCCTGCTC: ACTTCGTGGG AATAGTTCTC CGCTGAGATA: GCGCCTCTCG: 2160
CCTAAGTATT ATGTAAGTTG: ATTCCCTTC TTTTGTCTCT CTTGTTTGTG: CTATGCTTG 2220
ACCCAGTATC TCCCTCAAA: TGAAAGTTCT: CCGCTTGATT TTTCTGCTCT: GAAGGCAGGG 2280
TGAGTTCTCT CTTCAAGAA: GACTTCAAC CATTTAACGT GTTCTTAAG AGCCGTCAAT 2340
50 CAGCCTGGTT TTGGGATGC TATGAAAGAG AGAAGGAAA: TCATGCGCT: CAGTCTCTGG 2400
AGACAGAAGA GCCGTATCA GTGTCTCACT TGTGATTTT: ATCTGGAATA: GGAAGAACA: 2460
CCCCAGCAGA GCAAGCTCAG CTTTCTAGAG AAGGATATTT: CCAACTGCA: AACTTTGCTT 2520
TGAAAGTTT AGCCCTTTAA: GGAATGAAAT CATGTAGAAT TTTGAGCTTC: TAAAAACATT 2580
AAATCAGCT TATTAATACG GATAGAGAA: AGAAATCTGG: TGCCTGGGGG: TCCCTGTGTT: 2640
55 CACCCCTAGA GTTGTGTTTA: AAATTTTAA: TTGAAGCATG TGAAGTGTAC STGCAGAAAA: 2700
GTGGGAACAT GATAGTGTAT: GGCTTGGTGG ATTTTCACAA ACTGAACATA: CTTGTGTAAT 2760
CAGCATCTAG ACCCAGACCC AGAGCATCAC AAATATCCCC CATCTGGGC: TTTTCCAGA 2820
GGAGATGGGG GCTTCTGAAG ATGGACTTAC: CTGGGACCTG: CCCCCATGA: GCCAGGACGG: 2880
TCCCCCACA GTACGCTGT GCAAGGCCCC CGTGCCAGG GGTGGAGGAG: AATATGTGG 2940
60 TGTGGACAGG ATGGGAGAT GTGGCCTGAA: CAGGAGATT: TATTATATCT GGAGACCCCTG 3000
AGAGACCCCTG AGACCTGGGG CACCATGGCT: GGCCAGGTCA: GAAGCATCT: GACTGCAGAG: 3060
GTCCGTGAGC CCACACCCCTC TTCCCTGCCA: GCAAGTTGTC: TGGGCTCAT: CGAGGCCCC: 3120
TCCGCTGGA: GCCTTCTATG GACGTGATAT GCCTGTATCT: GTTTTAAAT TTTCTTCTC 3180
ACTTAGGGGA AGTGAATCG CTCAGAGATG AGATCCTTTA ATTGAACAG: AAGTGTAAAG 3240
75 GAATCTAGTG TCTTCTAAT GTGGTAAAT TCTCCATCAA CATCAGTTC AGCTGGCAGC: 3300
TGAACCTCAG AATCTCACTT ACAGCAGCG ACACGGGGT ACACCGATGG GTCCACTGG 3360
65 GTCTGGGGGC TCCCTGGAGC TCTCTCTGG GTGTGCTGG TTAGGAGTTG: AGTTGTTTGC 3420
TCCAGGGTTA TTTCTCTCTT CGAGTCAAG TCACACGAAT ACCTGCCTTC TCTGGCTTC 3480
CTGCTATACA CATATTACA TGGCGCTCAA GAAGTTAGGC TCATGGCAAC GTGTGTCTTT 3540
CTCTGGACAA CTGCCCCAG TTACAGTGAA ATGGAGAATT: TCAGTCTTCC AGCTCTGCCC 3600
70 AGGAAAGAAC TTCAGCTGAC TCCACGGGA: TCTGGAATC CACGACCAAT CCGCATCGGC 3660
TCTTATTAGC TCCCGCTCC ACAAGACACC TGTGCTTTGG AAATCCACCA CCAATCCGSA 3720
TGGGCTCTTA TTAGCTCCCC GCTCCACAAG ACACCTGTGA TCTGGAATC TACCACCAAT 3780
CCCGATGGC TCTTATTAGC TCCCGCTCC ACAAGACACC TGTGACATCC TCCAGGGCCA 3840
CAGGAGCACG TGCTGACCA CTTTCCCTTC CAGTTCCTGC ACAAAAGTG TCCAGAGGGC 3900
75 TGTTTGCAAA CACTAGTCCA CTTTGTAGCT TTTCAACCTC TGTCCAGGG AATCTAGGAG 3960
AGATGAGGCC CGTCAGATC AAGAGATGTC ATCCCCCAG GGTCTCCAG GCATTTCCAC 4020
ACTATTGGTG GCACCTGGAG GACATGCACC AAGGCTTGCC AGAGCCAACA GGAAGTGAGC 4080
CCAGAGCATG GCACATGAGC ATCACCCTGCT GATGGTGCC TGTGTGCTT GGTGCCAACA 4140
80 GGGGCATCCC GGCCGTAACC CCTCCAGAC: GGAAGCATGG GTTGGCCAC AGACCTGTGG 4200
GGTGTCTCTG TGAGTGGCCT CCAGATGTCT TGTGTCATG GCACAAGTG: GCCAGGGCTG 4260
GAGGGAGGTG GGAAACCTCA TCATCCGGTG GGCCCTGCCA ATCTTAACCC AGAACCTTTA 4320
GGTATTCTTG GCAGTAGCCA TGACATTGGA GCACCTTCTC CTCAGCCAG AGGCTGACCT 4380
GAGGGCACT GTCTCAGAT GACACCAACC AGGAGCACCC TAGGTGAGGG: GTGAGGGCCC 4440
CCTTATGTA ACCTCTTGCC TCTTCTTTC TCCCATCAGA GTGGTTGGAT GGAGCCATTG 4500
85 GCCTCCTTTT: CTTAGCGGG CCGTTCAACC TCTCTGACC ATGTTGTCTG: GCTGAGGAGC 4560
TACTAGAAA GCTGAGTGA GTCTCCTTC CAACAGGATG ATGATTTGC TCAATTTCTCA 4620
GGGCTGGAAT GAGCGGCTG GTCCCCAGA AAGCTGGAGT GGGGTACAGA GTTCAGTTT 4680
CCTCTCTGTT TACAGCTCCT TGACAGTCCC ACGCCCATCT GGAGTGGGAG CTGGGAGTTA 4740

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

Seq ID NO: 286 Protein sequence:
Protein Accession #: NP_570843.1

1 11 21 31 41 51
MPLKHYLLLL VGCQAWGAGL AYHGPCSECT CSRASQVECT GARIVAVPTP LPWNAMSLOI 60
LNTHITELNE SPFLNISALI ALRIEKNELS RITPGAFRNL GSLRYLSLAN NKLQVLPIGL 120
FQGLDSLESL LLSSNQLLQI QPAHFSQCSN LKELQLEHNE LEYIPDGAFT HLVLGLTKLNL 180
GKNSLTHISP RVFQHLNQL VLRLYENRLT DIPMGTDFGL VNLQELALQQ NQIGLLSPGL 240
FHNHNLQRL YLSSNHLISQL PPSIFMQLPQ LNRLLTFGNS LKELSLGIFG PMPNHLRELWL 300
YDNHISSELD NVFSNLRQLQ VLILSRNQIS FISPGAFNGL TELRELSLHT NALQDLQDGNV 360
FRMLANLQNI SLQNNRLRLQ PGNIFANVNG LMAIQLNQNG LENLPLGIFD HLGLKCLRLR 420
YDNFWRCDSD ILPLRLNWL L NQPRLGTDIV PVCFSPANVR GQSLIIINVN VAVPSVHVPE 480
VPSYPETPWY PDPSPYPTT SVSSTELTS FVEDYTDLIT IQVTDDRSVN GMTQAQSGLA 540
IAAIVIGIVA LACSLAACVG CCCCKRSQA VLMQMKAPNE C

Seq ID NO: 287 DNA sequence
Nucleic Acid Accession #: NM_002362
Coding sequence: 1..954

40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
ATGCTCTCTG AGCAGAAGAG TCAGCACTGC AAGCCTGAGG AAGCGTTGA GGCCCAAGAA 60
GAGGCCCTGG GCCTGGTGGG TGACAGGCT CCTACTACTG AGGAGCAGGA GGCTGCTGTC 120
TCCCTCTCT CTCTCTGCT CCTGGCACC CTGGAGGAAG TGCTGCTGTC TGAGTCAGCA 180
GGTCTCTCCC AGAGTOCTCA GGGAGCCTCT GCCTTACCCA CTACCATCAG CTTCACTTGC 240
TGGAGGCAAC CCAATGAGGG TTCCAGCAGC CAAGAAGAGG AGGGGCCAAG CACCTGCGCT 300
GACGAGAGT CCTTGTTCOG AGAAGCACTC AGTAACAAGG TGGATGAGTT GGCTCATTTT 360
CTGCTCCGCA AGTATCGAGC CAAGGAGCTG GTACACAAAG CAGAAATGCT GGAGAGAGTC 420
ATCAAAAATT ACAAGCGCTG CTTTCTGTG ATCTTCGGCA AAGCCTCGA GTCCCTGAA 480
ATGATCTTTG CANTGACGT GAAGGAAGTG GACCCCGCCA GCAACACCTA CACCCTGTG 540
ACCTGCCTGG GCCTTTCTTA TGATGGCCTG CTGGGAATA ATCAGATCTT TCCCAAGACA 600
GGCCTCTGTA TAATCTCTCT GGGCACAATT GCAATGGAGG GCGACAGCGC CTCGAGGAG 660
GAAATCTGGG AGGAGCTGGG TGTGATGGG GTGTATGATG GGAGGGAGCA CACTGTCTAT 720
GGGAGGCCCA GGAACCTGCT CACCAAGAT TGGGTGCAGG AAAACTACCT GGAGTACCGG 780
CAGGTACCCG CAGTAATCC TGGCGCTAT GAGTTCCTGT GGGGTCCAAG GGCTCTGGCT 840
GAAACCACT ATGTGAAGT CTTGGAGCAT GTGGTCAGG TCAATGCAAG AGTTGCGCAT 900
GCCTACCCAT CCTCGGTGA AGCAGCTTGT TTAGAGGAG AAGAGGAGT CTGA

Seq ID NO: 288 Protein sequence:
Protein Accession #: NP_002353.1

65
70
75
80
85

1 11 21 31 41 51
MSSEKQSHC KPEEGVEAQE EALGLVGAQA PTTEEQEAIV SSSSPLVPGT LEEVPAESA 60
GPPQSPQAS ALPTTISFTC WRQPNEGSSS QEEEGPSTSP DAESLFREAL SNKVDELAHP 120
LLRKYRAKEL VTKAEMLERV IKNYKRCFPV IPGKASESLK MIFGIDVKEV DPASNTYTLV 180
TCLGLSYDGL LGNNQIFPKT GLLIIVLGTI AMEGDSASEB EIWEELGVNG VYDGREHTVY 240
GEPRKLTQD WQENYVLEYR QVPGSNPARY EFLWGPRLA STSYVKVLEH VVRVNAVRI 300
AYPELRREAL LEEEBGV

Seq ID NO: 289 DNA sequence
Nucleic Acid Accession #: NM_002362
Coding sequence: 46..1344

80
85

1 11 21 31 41 51
CGGCGGCCGC GCCTGGTGTG GGTCCCACT GCTCTGGGG GCGCATGGA CGAGGCCGTG 60
GGCGACCTGA AGCAGGCGCT TCCCTGTGTG GCGAGTGGC CAACGGTCCA CGTGGAGGTG 120
CATCAGCGCG GCAGCAGCAC TGCAAGAAA GAAGACATAA ACCTGAGTGT TAGAAGGCTA 180
CTCAACAGAC ATAATATTGT GTTGGTGAT TACACATGGA CTGAGTTTGA TGAACCTTTT 240
TTGACCAGAA ATGTGCACTG TGTGTCTATT ATTGACACAG AATTAAAGGT TAAAGACTCA 300
CAGCCCATCG ATTTGAGTGC ATGCACTGTT GCACCTTCA TTTTCCAGCT GAATGAAGAT 360
GGCCCCAGCA GTGAAAATCT GGAGGAAGAG ACAGAAAACA TAATTGCAGC AAATCACTGG 420
GTTTCACTCG CAGCTGAATT CCATGGGCTT TGGGACAGCT TGATATAGA TGTGGAAGTC 480
AAATCCCATC TCCTCGATTA TGTGATGACA ACTTTACTGT TTTCAGACAA GAACGTCAAC 540

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

AGCAACCTCA TCACCTGGAA CCGGGTGGT CTGCTCCAGG GTCCTCCTGG CACTGGAAAA 600
ACATCCCTGT GTAAAGCGTT AGCCACAGAAA TTGACAATTA GACTTTCAAG CAGGTACCGA 650
TATGGGCAAT TAATTGAAT AAACAGCCAC AGCCTCTTTT CTAAGTGGTT TTCGAAAGT 720
GGCAAGCTGG TAAACCAAGT GTTTCAGAAG ATTCAAGATT TGATTGATGA TAAAGACGCC 780
CTGGTGTTCG TGCTGATTGA TGAGGTGGAG AGTCTCACAG CGCCCCGAAA TGCTTCGAGG 840
GCGGGCACCG AGCCATCAGA TGCCATCCCG GTGGTCAATG CTGTCTTGAC CCAAATTGAT 900
CAGATTAAAA GGCATTCCAA TGTGTGATT CTGACCATT CTAACATCAC CGAGAAGATC 960
GACGTGGCCT TCGTGGACAG GGCTGACATC AAGCAGTACA TTGGGCCACC CTCTGCAGCA 1020
GCCATCTTCA AAATCTACCT CTCTGTGTTG GAAGAATGA TGAAGTGTCA GATCATATAC 1080
CCTCGCCAGC AGCTGTCTGAC CCTCCGAGAG CTAGAGATGA TTGGCTTCAT TGAACAACAC 1140
GTGTCAAAAT TGAGCCTTCT TTTGAATGAC ATTTCAAGGA AGAGCGAGGG CCTCAGCCGC 1200
CGGTCTCTGA GAAAACTCCC TTCTCTGGCT CATGCGCTGT ATGTCCAGGC CCCCACCGTC 1260
ACCATAGAGG GGTTCCTCCA GGCCCTGTCT CTGGCAGTGG ACAAGCAGTT TGAAGAGAGA 1320
AAGAAGCTTG CAGCTTACAT CTGATCCTGG GCTTCCCCAT CTGGTGTCTT TCCCATGGAG 1380
AACACACAC CAGTAAGTGA GGTGGCCCA CACAGCGCTC TCCAGGGGAA TCCTTCTGCT 1440
AAACCAACG TTAATCTAGC TGCAAGCTAG AAAGCCACCA AGGCAGGGCT TTGTAAAAAG 1500
AAGTGTATTC TATTATGTT GTTTTAAAT GCATACTAG AGACAAACAT CTGTCTATT 1560
TCACTGTTTG TAAAGATAA TTCAGATTGT TTGCTCCTT GTGAAGAACC ATCGAAACCT 1620
GTTTGTTCCT AGCCCCCCCC CAGTGGATGG GATGCATAAT GCCAGCAAGT TTTGTTTAA 1680
AGCAAAAAAG GAAGATTAAAT GCAGGTGTTA TAGAAGCCAG AAGAGAACT GTGTCACCT 1740
AAGAAGCAT ATAATCATAG CATTAAAAAT GCACACATTA CTCCAGGTGG AAGGTGGCAA 1800
TTGCTTCTG ATATCAGCTC GTTTGATTGA GTGCAAAAAT GTTTTCAAGA CTATTTAATG 1860
GATGTAAAAA AGCCTATTTC TACATTATAC CAACTGAGAA AAAAATGGTC GGTAAAGTGT 1920
TCCTTCATAA TAAATTAATCA AGACATGGTC CCAATTGCGA GAAAGTGCAT GACTCTGAGT 1980
TTTCCAGGGA AACACATGCT GGACATCCCT TGTAACCCGG TATGGCGGCC CCTGCATTGC 2040
TGGGATGTTT CTGCCCCAGG TTTTGTGTT GCAATAACGT TATCACATT CTAAATGAGGA 2100
TTCACATTAA TATAATATAA AATAAATAGG TCAGTTACTG GTCTCTTCT CCGCAATGTT 2160
ATGTTTGTCT TTTATCTCAC AGTAAATAA ATATAATTAA AAA

Seq ID NO: 290 Protein sequence:
Protein Accession #: NP_004228

1 11 21 31 41 51
MDEAVGLKQ ALPCVAESPT VHVEVHQRGS STAKKEDINL SVRKLNRHN IVFGDYTWE 60
FDEPFLTRNV QSVSIIDTEL KVKSQPIDL SACTVALHIP QLNEDGPSSE NLEEETENII 120
AANHWWLPAA EFHGLWDSL VYDEVKSHLL DYVMTLLFS DKNVNSNLIT WNRVVLHGP 180
PGTKTSLCK ALAQKLIRL SSRVRYQLI EINHSLFSK WFSSEGLVT KMFQIKDLI 240
DDKDALVFLV IDEVESLTA RACRAGTEP SDAIRVNVAV LTQIDQIKRH SNVILITSN 300
ITEKIDVAFV DRADIKQYIG PPSAAIFKI YLSCEELMK CQIIYPROQL LTLRELEMIG 360
FIENNVSRLS LLLNDISRKS EGLSGRVLK LPFLAHALYV QAPVTITIEGP LQALSLAVDK 420
QFEERKKLAA YI

Seq ID NO: 291 DNA sequence
Nucleic Acid Accession #: NM_002658.1
Coding sequence: 77-1372

50
55
60
65
70
75
80
85

1 11 21 31 41 51
GTCCCGCAG GCGCGTCGCG CCTCCTGCC GCAGGCCACC GAGGCCGCG CCGTCTAGCG 60
CCCGGACCTC GCCACCATGA GAGCCCTGCT GCGCGGCTG CTCTCTGCG TCCTGTGCT 120
GAGCGACTCC AAGGCGACA ATGAATCTCA TCAAGTCCCA TCGAAGTGTG ACTGTCTAAA 180
TGAGGAACA TTGTGTGCCA ACAAGTACTT CTCCAAACAT CACTGGTGC ACGGCCAAA 240
GAAATTCGGA GCGCAGCACT GTGAAATAGA TAAGTCAAAA ACCTGCTATG AGGGGAATGG 300
TCACCTTTAC CGAGGAAAGG CCAGCACTGA CACCATGGCG CGGCCCTGCC TCCCTGGAA 360
CTCTGCCACT GTCCCTCAGC AAACGTACCA TGCCCAAGA TCTGATGCTC TTCAGCTGGG 420
CCTGGGGGAA CATATAACT GCAGGAACCC AGACAACCGG AGGCGACCTC GGTGCTATGT 480
CGAGGTGGGC CTAAAGCGCG TTGTCCAAGA GTGCATGGTG CATGACTGCG CAGATGGAAA 540
AAAGCCCTCC TCTCTCCAG AAGAATTAAA ATTTCAAGT GGCCTAAAGA CTCTGAGGCC 600
CGCTTTAAG ATTATTGGGG GAGAATTCAC CACCATCGAG AACACAGCCT GGTTCGCGC 660
CATCTACAG AGGCACCGGG GGGCTCTGT CACCTACGTG TGTGGAGGCA GCCTCATCAG 720
CCCTGTCTGG GTGATCAGCG CCACACACTG CTTCATTGAT TACCCAAAGA AGGAGGACTA 780
CATGCTCTAC CTGGGTGCTC CAAGGCTTAA CTCCAACAG CAAGGGGAGA TGAAGTTTGA 840
GGTGGAAAAC CTCATCTAC ACAGGACTA CAGCGCTGAC ACGCTTGCTC ACCCAACGA 900
CATTGCTCTG CTGAAGATCC GTTCCAAGGA GGGCAGGTGT GCGCAGCCAT CCGGACTAT 960
ACAGACCATC TGCCCTGCCCT CGATGTATAA CGATCCCCAG TTTGGCACA GCTGTGAGAT 1020
CACTGGCTTT GGAAAGAGA ATTCTACCGA CTATCTCTAT CCGGAGCAGC TGAAAATGAC 1080
TGTGTGAAG CTGATTTCCC ACCGGGAGTG TCAGCAGCCC CACTACTACG GCTCTGAAGT 1140
CACCACCAA ATGCTATGTG CTGCTGACCC CCAATGGAAA ACAGATTCTT GCCAGGAGA 1200
CTCAGGGGGA CCCCTCGTCT GTTCCTCCA AGGCCGCTG ACTTTGACTG GAATTGTGAG 1260
CTGGGGCCGT GGATGTGCC TGAAGGACAA GCCAGGCTC TACACGAGAG TCTCACACTT 1320
CTTACCTCGG ATCCGCTGAC ACACCAAGGA AGAGAATGGC CTGGCCCTCT GAGGGTCCCC 1380
AGGAGGAAA CCGGCACCA CCGCTTCTT GCTGGTTGTC ATTTTGCAG TAGAGTCATC 1440
TCCATCAGCT GTAAGAAGAG ACTGGGAAGA TAGGCTCTGC ACAGATGGAT TTGCCTGTGG 1500
CACCACCAAG GTGAACGACA ATAGCTTTAC CCTCACGGAT AGGCCTGGGT CTGGCTGCC 1560
CAGACCTCTT GGCCAGGATG GAGGGGTGTT CCGTACTCAA CATGTTACTG ACCAGCAACT 1620
TGTCTTTTTC TGGATCTAAG CTGTCAGGAG TTAAGAAAGG CAGGGCATCT CCTGTGCATG 1680
GGCTCGAAGG GAGAGCCAGC TCCCCGACC GGTGGGCATT TGTGAGGCC ATGGTTGAGA 1740
AATGAATAAT TTCCCAATTA GGAAGTGTAA GCAGCTGAGG TCTCTTGAGG GAGCTTAGCC 1800
AATGTGGGAG CAGCGGTTTG GGGAGCAGAG ACACATAAGA CTTCAGGGCA GGCCTCTGAT 1860
ATTCCATGAA TGTATCAGGA AATATATATG TGTGTGATG TTTGCACACT TGTGTGTG 1920
GCTGTGAGTG TAAGTGTGAG TAAGAGCTGG TGTCTGATT TTAAGTCTAA ATATTCTCTT 1980
AAACTGTGTG GACTGTGATG CCACACAGAG TGGTCTTCT GGAGAGGTTA TAGGTCATC 2040
CTGGGGCTC TTTGGTCCCC CAGCTGACAG TGCCCTGGGA TGTACTTATT CTGCAGCATG 2100
ACCTGTGACC AGCACTGTCT CAGTTTCACT TTCACATAGA TGTCCCTTTC TTGGCCAGTT 2160
ATCCCTTCTT TTTAGCCTAG TTCATCCAAT CCTCACTGGG TGGGGTGAAG ACCACTCTCT 2220
ACACTGAATA TTTATATTTC ACTATTTTTA TTTATATTTT TGAATTTTA AATAAAGTG 2280

WO 02/086443

PCT/US02/12476

ATCAATAAAA TGTGATTTTT CTGA

Seq ID NO: 292 Protein sequence:
Protein Accession #: NP_002649.1

5
10
15

1	11	21	31	41	51	
MRALLARLLL	CVLVVSDDSKG	SNELHOVPSN	CDCLNGGTCV	SNKYFSNIHW	CNCPKFGGQ	60
HCEIDKSKTC	YEGNGHFYRG	KASDTMGRP	CLPWSATVL	QQTTHAHRSD	ALQLGLGKH	120
YCRNPDNRNR	PWCYVQVGLK	PLVQECMVHD	CADGKXPSSP	PEELKFQCGQ	KTLLRPFKII	180
GGEFTTIENQ	PWFAAIYRRH	RGGSVTYVCG	GSLISPCWVI	SATHCIDYD	KKEDIYVYL	240
RSRLNSNTQG	EMKFEVENLI	LHKDYSADTL	AHNDIALLK	IRSKGRCQAQ	PSRTIQITCL	300
PSMYNDPQFG	TSCEITQFGK	ENSTDYLYPE	QLKMTVVKLI	SHRECQPHY	YGSEVTTKML	360
CAADPQWKTD	SCQDSDGGPL	VCSLQGRMTL	TGIVSWGRGC	ALKDKPGVYT	RVSHFLPWIR	420
SHTKEENGLA	L					

Seq ID NO: 293 DNA sequence
Nucleic Acid Accession #: NM_001498
Coding sequence: 93..2006

20
25
30
35
40
45
50
55
60
65

1	11	21	31	41	51	
GGCAGGAGGC	TGAGTGTCCG	TCTCGCGCCC	GGAAGCGGGC	GACCGCCGTC	AGCCCGGAGG	60
AGGAGGAGGA	GGAGGAGGAG	GAGGGGGCGG	CCATGGGGCT	GCTGTCCAG	GGCTGCGCG	120
TGAGCTGGGA	GGAAACCAAG	CGCCATGCCG	ACCACTGGCG	GGCGACGGG	ATCTCCAGT	180
TCCTGCACAT	CTACACGCC	GTCAAGGACC	GGCACAAGGA	CGTTCTCAAG	TGGGGCGATG	240
AGGTGGAATA	CATGTTGGTA	TCTTTTGATC	ATGAAAATAA	AAAAGTCCGG	TGTTCTCTGT	300
CTGGGGAGAA	AGTTCCTTGA	ACTCTGCAAG	AGAAGGGGGA	AAGGACAAAC	CCAAACCATC	360
CTACCCCTTG	GAGACGAGAG	TATGGGAGTT	ACATGATTGA	AGGGACACCA	GGACAGCCCT	420
ACGGAGGAAC	AATGTCCGAG	TTCATACAG	TTGAGGCCAA	CATGCGAAAA	CGCCGGAAGG	480
AGGCTACTTC	TATATTAGAA	GAAATCAGG	CTCTTTGCAC	AATAACTTCA	TTTCCAGAT	540
TAGGCTGTCC	TGGGTTCAAC	CTGCGCGAGG	TCAAACCCAA	CCCAGTGGAA	GGAGGAGCTT	600
CCAAGTCCCT	CTTCTTTCCA	GATGAAGCAA	TAAACAAGCA	CCCTCGCTTC	AGTACCTTAA	660
CAAGAAATAT	CCGACATAGG	AGAGGAGAAA	AGGTTGTGAT	CAATGTACCA	ATATTTAAGG	720
ACAGAAATAC	ACCATCTCCA	TTTATAGAAA	CATTTACTGA	GGATGATGAA	GCTTCAAGGG	780
CTTCTAAGCC	GGATCATATT	TACATGGATG	CCATGGGATT	TGGAATGGGC	AATTGCTGTC	840
TCCAGGTGAC	TTTCCAAGCC	TGCAGTATAT	CTGAGGCCAG	ATACCTTTAT	GATCAGTTGG	900
CTACTATCTG	TCCAATTGTT	ATGGCTTTGA	GTGCTGCATC	TCCCTTTTAC	CGAGGCTATG	960
TGTGACACAT	TGATTGTGCG	TGGGAGTGA	TTTCTGCATC	TGTAGATGAT	AGAACTCGGG	1020
AGGAGCGAGG	ACTGGAGCCA	TTGAAGAACA	ATAACTATAG	GATCAGTAAG	TCCCGATATG	1080
ACTCAATAGA	CAGTATTTTA	TCTAAGTGTG	GTGAGAAATA	TAATGACATC	GACTTGACGA	1140
TAGATAAAGA	GATCTACGAA	CAGCTGTTGC	AGGAAGGCAT	TGATCATCTC	CTGGCCGAGC	1200
ATGTTGCTCA	TCTCTTTATT	AGAGACCCAC	TGACACTGTT	TGAAGAGAAA	ATACACCTGG	1260
ATGATGCTAA	TGAGTCTGAC	CATTTTGAGA	ATATTGAGTC	CACAAATTGG	CAGACAATGA	1320
GATTTAAGCC	CCCTCTCCCA	AACCTCAGACA	TTGGATGGAG	AGTAGAATTT	CGACCCATGG	1380
AGGTGCAATC	AACAGACTTT	GAGAACTCTG	CCTATGTGGT	GTTTGTGGTA	CTGCTCACCA	1440
GAGTGATCCT	TTCTTCAAAA	TTGGATTTC	TCATTCCACT	GTCAAAGGTT	GATGAGAAAC	1500
TGAAGGTAGC	ACAGAAAAGA	GATGCTGTCT	TGCAGGGAAT	GTTTATTTC	AGGAAAGATA	1560
TTTGCAAAAG	TGGCAATGCA	GTGGTGGATG	GTTGTGGCAA	GGCCCAAGAC	AGCAGCGAGC	1620
TOGCTGCAGA	GGAGTACACC	CTCATGAGCA	TAGACACCAT	CATCAATGGG	AAGGAAGGTG	1680
TGTTTCTCGG	ACTGATCCCA	ATTCTGAAC	CTTACCTTGA	AAACATGGAA	GTGGATGTGG	1740
ACACCAAGTG	TAGTATTCTG	AACTACCTAA	AGCTAATTAA	GAAGAGAGCA	CTCGGAGAAC	1800
TAATGACAGT	TGCCAGATGG	ATGAGGGAGT	TTATGCAAAA	CCATCCTGAC	TACAGCAAG	1860
ACAGTGTGAT	AACGTATGAA	ATGAATTATA	GCCTTATTTT	GAAGTGAAC	CAAAATTGCA	1920
ATGAATTATG	TGAATGCCCA	GAGTTACTTG	GATCAGCATT	TAGGAAAGTA	AAATATAGTG	1980
GAAGTAAATC	TGACTCATCC	AACAGACAT	TCTACAGAAA	GAAAAATGCA	TTATTGACGA	2040
ACTGGCTACA	GTTCAATGCG	TCTCAGCCCG	TGTGTATAAT	ATGAAGACCA	AATGATAGAA	2100
CTGTACTGTT	TTCTGGGCCA	GTGAGCCAGA	AATTGATTAA	GGCTTTCTTT	GGTAGGTAAA	2160
TCTAGAGTTT	ATACAGTGTA	CATGTACATA	GTAAGTATT	TTTGATTAA	AATGTATTTT	2220
AATAACATAT	CTAAAGTCAT	CATGAACCTG	CTGTACATT	TTTAAATTCT	TACTCTGGAG	2280
CAACCTACTG	TCTAAGCAGT	TTTGTAAATG	TACTGGTAAT	TGTACAATAC	TTGCAITCCA	2340
GAGTTAAAT	GTTTACTGTA	AATTTTGTG	CTTTTAAAGA	CTACCTGGGA	CCTGATTAT	2400
TGAAATTTT	CTCTTTAAAA	ACATTTCTC	TGTTAAATTT	TCCTTTGTCA	TTCTCTTTGT	2460
TGTCTACATT	AAATCACTTG	AATCCATTGA	AAGTGCTTCA	AGGGTAATCT	TGGGTTTCTA	2520
GCACCTTATC	TATGATGTTT	CTTTTGCAAT	TGGAATAATC	ACTTGGTCAC	CTTGCCCCAA	2580
GCTTTCCCT	CTGAATAAAT	ACCCATTGAA	CTCTGAAAA	AAAAAATAAA	AAAA	

Seq ID NO: 294 Protein sequence:
Protein Accession #: NP_001489

70
75
80
85

1	11	21	31	41	51	
MGLLSQGSPL	SWEETKRHAD	HVRRHGILQF	LHIYHAVKDR	HKDLVKWDE	VEYMLVSFDH	60
ENKKVRLVLS	GEKLETLQE	KGERTNPNHP	TLWRPEYGSY	MEBTPGQPY	GGTMSFNTV	120
EANMRKRRKE	ATSILENQA	LCTITSFPRL	GCPGFTLPEV	KPNPVEGGAS	KSLFFPDEAI	180
NKHPRFSTLT	RNIRHRRGEK	VVINVPFKD	KNTSPPIET	FTEDDEASRA	SKPDHIYMDA	240
MFGMGNCCL	QVTFOACIS	EARYLYDQLA	TICPIVMALS	AASPFYRGYV	SDIDCRWGI	300
SASVDDRTRE	ERGLEPLNN	NYRISKSRYD	SIDSYLSKOG	EKYNDIDLTI	DKEIYEQLLQ	360
EGIDHLLAQH	VAHLEIRDL	TLFEKIHLD	DANESDHFEV	IQSTNWQTM	PKPPPPNSDI	420
GWRVEFRPME	VQLTDFENSA	YVVFVLLTR	VILSYKLDL	IPLSKVDENM	KVAQKRDAVL	480
QGMFYFRDI	CKGNAVVDG	CGKAQNSTEL	AAEYTLMSI	DTIINGKEGV	FPGLIPILNS	540
YLENMEVDVD	TRCSILNYLK	LKKRASGEL	MTVARWMREP	IANHPDYKQD	SVITDEMYS	600
LILKCNQIAN	ELCECPELLG	SAFRKVKYSG	SKTDSSN			

WO 02/086443

PCT/US02/12476

Seq ID NO: 295 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-816

```
5 1 11 21 31 41 51
AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCTCCCCA TGCACATCAC TCCCAGGTGC 120
CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
10 GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CCGTCTCTCC TCCAGCAAGG 240
GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAAGCTT 420
ATGACAGGAC ATGCTATTCC ACCCAGCCAA TTGGATTCTC AGATTGATGA CTTCACTGGT 480
15 TTCAGCAAAAG ATAGGATGAT GCAGAAACCT GGTAGCAATG CACCTGTGGG AGGAAACGTT 540
ACCAGCAGTT TCTCTGGAGA TGACCTAGAA TGCAGAGAAA CAGCCTCCTC TCCCAAAGC 600
CAACGAGAAA TTAATGCTGA TATAAAACGT AAATAGTGA AGGAACTCCG ATGCGTTGGA 660
CAAAAATATG AAAAATCTTT CGAAATGCTT GAAGGAGTGC AAGGACCTAC TGCAGTCAGG 720
AAGCGATTTT TTGAATCCAT CATCAAGGAA GCAGCAAGAT GTATGAGACG AGACTTTGTT 780
20 AAGCACCTTA AGAAGAACTC GAAACGTATG ATTTGAGAACT ACTTGTCCCT GGAGGATTAT 840
CACACCCCAA ATGCATAATC TCGTTAATGA TTGAGGAGAG AAAAGGATCA GATTGCTGTT 900
TTCTACAATG GAGCAGGATA TTGCTGAAGT CTCTGGCAT ATGTTACCGA ATCAAAATAGC 960
CTTCCAGAGG CTAAGAAATT TCTGTTAGTA AAAGATGTTT TTTTCCCAA AGCATTITAT 1020
TTGAAAGGAT AACTTGTGTT TTGTTTATTT TGTATTCCCA CCTGTGCTGG TAGATATTAT 1080
25 TAACCCATTA GGTAAATACT ATTACAGTCG TGGTTTCTGC A
```

Seq ID NO: 296 Protein sequence:
Protein Accession #: Eos sequence

```
30 1 11 21 31 41 51
MTDKTEKVV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKEKKLMT 60
GHAIPPSQLD SQIDDFTGFS KDRMMQKPGS NAPVGGNVTS SFGDDLECR ETASSPKSQR 120
EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR PFESIIKEAA RCMRRDFVKH 180
35 LKKKLKRM
```

Seq ID NO: 297 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-815

```
40 1 11 21 31 41 51
AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCTCCCCA TGCACATCAC TCCCAGGTGC 120
CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
45 GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CCGTCTCTCC TCCAGCAAGG 240
GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAAGCTT 420
ATGACAGGAC ATGCTATTCC ACCCAGCCAA TTGGATTCTC AGATTGATGA CTTCACTGGT 480
TTCAGCAAAAG ATAGGATGAT GCAGAAACCT GGTAGCAATG CACCTGTGGG AGGAAACGTT 540
ACCAGCAGTT TCTCTGGAGA TGACCTAGAA TGCAGAGAAA CAGCCTCCTC TCCCAAAGC 600
CAACGAGAAA TTAATGCTGA TATAAAACGT AAATAGTGA AGGAACTCCG ATGCGTTGGA 660
55 CAAAATATG AAAAATCTTT CGAAATGCTT GAAGGAGTGC AAGGACCTAC TGCAGTCAGG 720
AAACGATTTT TTGAATCCAT CATCAAGGAA GCAGCAAGAT GTATGAGACG AGACTTTGTT 780
AAGCACCTTA AGAAGAACTC GAAACGTATG ATTTGAGAACT ACTTGTCCCT GGAGGATTAT 840
CACACCCCAA ATGCATAATC TCATTAATGA TTGAGGAGAG AAAAGGATCA GATTGCTGTT 900
TTCTACAATG GAGCAGGATA TTGCTGAAGT CTCTGGCAT ATGTTACCGA ATCAACTGGC 960
60 CTTCCAGAGG CTAAGAAATT TCTGTTAGTA AAAGATGTTT TTTTCCCAA AGCATTITAT 1020
TTGAAAGGAT AACTTGTGTT TTGTTTATTT TGTATTCCCA CCTGTGCTGG TAGATATTAT 1080
TAACCCATTA GGTAAATACT ATTACAGTCG TGGTTTCTGC A
```

Seq ID NO: 298 Protein sequence:
Protein Accession #: Eos sequence

```
65 1 11 21 31 41 51
MTDKTEKVV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKEKKLMT 60
GHAIPPSQLD SQIDDFTGFS KDRMMQKPGS NAPVGGNVTS SFGDDLECR ETASSPKSQQ 120
EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR PFESIIKEAA RCMRRDFVKH 180
70 LKKKLKRM
```

Seq ID NO: 299 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-815

```
75 1 11 21 31 41 51
AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCTCCCCA TGCACATCAC TCCCAGGTGC 120
CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
80 GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CCGTCTCTCC TCCAGCAAGG 240
GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGC AAAGAAGCTT 420
85 TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCCTGTTT 480
```


WO 02/086443

PCT/US02/12476

TCAGCAAGA TAGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAACGTTA 540
CCAGCAGTTT CTCTGGAGAT GACCTAGAAT GCAGAGAAAC AGCCTCCTCT CCCAAAGGCC 600
AACAAGAAAT TAATGCTGAT ATAAACCGTA AATTAGTGAA GGAACCTCGA TCGTGTGGAC 660
AAAAATATGA AAAATCTCTC GAAATGCTTG AAGGAGTGCA AGGACCTACT GCAGTCAGGA 720
AACGATTTTT TGAATCCATC ATCAAGGAAG CAGCAAGATG TATGAGACGA GACTTTGTGA 780
AGCACCTTAA GAAGAAACCTG AAACGTATGA TTTGAGAATA CTGTGCTCTG GAGGATTATC 840
ACACCCCAAA TGCAATAATCT CATTAAATGAT TGAGGAGAGA AAAGGATCAG ATTGCTGTGT 900
TCTACAATGG AGCAGGATAT TGCTGAAGTC TCCTGGCATA TGTACCGGAA TCAACTGGCC 960
TTCCAGAGGC TAAGAAATTT CTGTAGTAA AAGATGTTCT TTTTCCCAAA GCGTTTATT 1020
TGAAAGGATA ACTTGTGTTT TGGTTATTTT GTATTCCAC CTGTGCTGGT AGATATTATT 1080
AACCCATTAG GTAAATACTA TTACAGTCGT GGTTCCTGCA

Seq ID NO: 300 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
MTDKTEKVV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKAKKLMT 60
GHAIPPSQLD SQIDDFTFPS KDRMMQKPGS NAPVGGNVTS SPSGDDLECR ETASSPKSQQ 120
EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGTAVRKR FFSIIKEAA RCMRRDFVKH 180
LKKLKRMI

Seq ID NO: 301 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-812

1 11 21 31 41 51
AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCTCCCCA TGCACATCAC TCCAGGTGC 120
CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
GGGAGGGGCC ACAACTTCAC TGCCATTTTG TGAGGTGCGG CGTCTCTCC TCCAGCAAGG 240
GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAAGCT 300
CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAGCTTA 420
TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCCTGGTT 480
TCAGCAAGA TGCGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAACGTTA 540
CCAGCAATTT CTCTGGAGAT GACCTAGAAT GCAGAGGAAT AGCCTCCTCT CCCAAAGGCC 600
AACAGAAAT TAATGCTGAT ATAAATGTC AAGTAGTGAA GGAAATCGA TGCCTTGGAC 660
AATATGAAA AATCTTCGAA ATGCTTGAAG GAGTGCAAGG ACCTACTGCA GTCAGGAAAC 720
GATTTTTTGA ATCCATCATC AAGGAAGCAG CAAGATGTAT GAGACGAGAC TTTGTTAAGC 780
ACCTTAAGAA GAAACTGAAA CGTATGATTT GAGAATACTT GTCCCTGGAG GATTATCACA 840
CCCCAATGCA ATAACTCTAT TAATGATTGA GGAGAGAAAA GGATCAGATT GCTGTTTCT 900
ACATGGAGC AGGATATTGC TGAAGTCTCC TGGCATATGT TACCGAATCA ACTGGCCTTC 960
CAGAGCTTAA GAAATTTCTG TTAGTAAAG ATGTTCTTTT TCCCAAAGCG TTTTATTGTA 1020
AAGGATAACT TGTGTTTGG TTATTTTGTA TTCCACCTG TGCTGGTAGA TATTATTAAAC 1080
CCATTAGGTA AATACTATTA CAGTCGTGGT TTCTGCA

Seq ID NO: 302 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
MTDKTEKVV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKEKKLMT 60
GHAIPPSQLD SQIDDFTFPS KDRMMQKPGS NAPVGGNVTS NFSGDDLECR GLASSPKSQQ 120
EINADIKQV VKELRCVGQK YEKIFEMLEG VQGTAVRKR FFSIIKEAA RCMRRDFVKH 180
KKKLKRMI

Seq ID NO: 303 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-815

1 11 21 31 41 51
AGTGTTCGGC TGGGACAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTTCC ATCCCCCTTG 60
GGCCAAACAG GATCGGTGCT TCTGGTGAGA CGTCTCCCCA TGCACATCAC TCCAGATGC 120
CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
GGGAGGGGCC ACAACTTCAC TGCCATTTTG TGAGGTGCGG CGTCTCTCC TCCAGCAAGG 240
GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAAGCT 300
CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGC AAAGAGCTTA 420
TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCCTGGTT 480
TCAGCAAGA TAGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAACGTTA 540
CCAGCAGTTT CTCTGGAGAT GACCTAGAAT GCAGAGAAAC AGCCTCCTCT CCCAAAGGCC 600
AACAGAAAT TAATGCTGAT ATAAACCGTA AATTAGTGAA GGAACCTCGA TCGTGTGGAC 660
AAAAATATGA AAAATCTCTC GAAATGCTTG AAGGAGTGCA AGGACCTACT GCAGTCAGGA 720
AACGATTTTT TGAATCCATC ATCAAGGAAG CAGCAAGATG TATGAGACGA GACTTTGTGA 780
AGCACCTTAA GAAGAAACCTG AAACGTATGA TTTGAGAATA CTGTGCTCTG GAGGATTATC 840
ACACCCCAAA TGCAATAATCT CGTTAATGAT TGAGGAGAGA AAAGGATCAG ATTGCTGTGT 900
TCTACAATGG AGCAGGATAT TGCTGAAGTC TCCTGGCATA TGTACCGGAA TCAACTGGCC 960
TTCCAGAGGC TAAGAAATTT CTGTAGTAA AAGATGTTCT TTTTCCCAAA GCGTTTATT 1020
TGAAAGGATA ACTTGTGTTT TGGTTATTTT GTATTCCAC CTGTGCTGGT AGATATTATT 1080
AACCCATTAG GTAAATACTA TTACAGTCGT GGTTCCTGCA

Seq ID NO: 304 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
5 MIDTEKVVAV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIIAGS AMSKAKKLMT 60
GHAIPPSQLD SQIDDFTFPS KDRMMQKPGS NAPVGGNVTS SPSGDDLECR BTASSPKSQQ 120
EINADIKRKL VKELRCVGQK YEKIFEMLEB VQGPVAVRKR FFEIIKEAA RCMRRDFVKH 180
LKKKLKRM

10 Seq ID NO: 305 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 87-689

1 11 21 31 41 51
15 CGTGGAGGCA GCTAGGCGCA GGCTGGGGAG GCTGAGCGG GCGTCTGTGC CCTGCGCTGC 60
CCAGACTAGC GAACAATACA GTCAGGATGG CTAAGAGTGA CCGCAAGAAA CCAAGGGGCA 120
AGATGTCGCG TTATGCTCTC TTTGTGCAGA CATGCAGAGA AGAACATAAG AAGAAAAACC 180
CAGAGGTCCC TGTCATTTT GCGGAATTTT CCAAGAAAGT CTCTGAGAGG TGGAAGACGA 240
TGTCGGGGAA AGAGAAATCT AAATTTGATG AAATGGCAAA GGCAGATAAA GTGCGCTATG 300
ATCGGGAAT GAAGGATTAT GGACAGCTA AGGGAGGCAA GAAGAAGAAG GATCCTAATG 360
CTCCCAAAAG GCCACCGTCT GATTTCTTCC TGTTCCTGTC AGAATTCGCG CCCAAGATCA 420
AATCCACAAA CCGCGGCATC TCTATTGGAG ACGTGGCAAA AAAGCTGGGT GAGATGTGGA 480
ATAATTATAA TTGACATGAA AAGCAGCCTT ACATCACTAA GCGCGCAAG CTGAAGGAGA 540
AGTATGAGAA GGAATGTGCT GACTATAAGT CGAAAGGAAA GTTTGATGGT GCAAAAGGTC 600
CTGCTAAAGT TGCCCGGAAA AAGTGGGAAG AGGAAGATGA AGAAGAGGAG GAGGAAGAAG 660
25 AGGAGGAGGA GGAGGAGGAG GATGAATAAA GAAACTGTTT ATCTGTCTCC TTGTGAATAC 720
TTAGAGTAGG GGAGGCGCGT AATTGACACA TCTCTTATTT GAGAAGTGTG TGTGCGCTGC 780
ATTAGTTTAA ATTACAAAAT TTGATCACGA TCATATTGTA GTCTCTCAA GTGCTCTAGA 840
AATGTGCTG GGTTCATCAT AAGTGGCCAT GGGTGTCTGG AGCACCTCGA AACTGTATCA 900
AAGTTGATA TAATTTCAAA CATTTTAAAT ATGAAAAGGC ACTCTGCTGT TCTCTCACT 960
30 CTGTGCACTT TGCTGTGGT GTGACAAGGC ATTTAAAGAT GTTCTGCGCA TTTTCTTTTT 1020
ATTGTAAAGG TGGTGGTAAC TATGTTTATT GGCTAGAAAT CCTGAGTTTT CAACGTATTA 1080
TATCTATAGT TTGTAAAGAA AACAAACAAA CCGAGACAAA CCCTTGATGC TCCTTGCTAG 1140
CGGTTGAGCG TGTTGGGAAAG ATGCTTTTGG GGAGAGGCTG TAGCTCAGGG CGTGCACTGT 1200
GAGGCTGGAC CTGTTGACTC TGCAGGGGCG ATCCATTTAG CTTGAGTTG TCTGTGTTCT 1260
35 GTATATAGTG ACATAGCATT CTGCTGCCAT CTTAGCTGTG GACAAAGGGG GGTCACTGG 1320
CATGAGAATA TTTTCTTTT TAAGTGGCGT AGTTTTTAAA CTGTTGTGTT TTAACAAAC 1380
TATAGAACTC TTCAITGTGA GCAAGCAAAA GAGTCACTGC ATCAATGAAA GTTCAAGAAC 1440
CTCCTGTACT TAAACACGAT TCGCAACGTT CTGTATTTT TTTTGTATGT TTAGAATGCT 1500
GAAATGTTT TGAAGTTAAA TAAACAGTAT TACATTTTAA AAACCTCTCT CTATTATAAC 1560
40 AGTCAATTTT TGACTCACAG CAGTGAACAA ACCCCCACTC CATTTGATTT GGAGACTGGC 1620
CTCCCTATAA ATGTGTGAGC TTCTTTTATT ACTCAGTGGC CAGCTCACTT AGGGCTGAGA 1680
TGAAGAGGAG GGCTACTTGA AGCTACTGTG TGATTTTGTG TGTGTCTGAG TGGCATTCAG 1740
ATGAAGTCTG GAGGAGTTAG GAGAAACGCA TAGGCAAGGT TCAGCAGCCT TCCAAGGTAT 1800
AGGAAGGTGG GTGATTGACT CTGAGGCTAT CTAGGTTTAA CTTTGTGCCC ACCTCCACCC 1860
45 CCTATTTTGT GGGGCCAAAT GCATTGCTAA ACAGCAATTT CAGAGTGTAT GGTGTGTCAA 1920
AAATTAAGGC CTTATGTTT TTCTCTTTCA CCCTACCCCG CCGTGTCTCT GGCACATATC 1980
ACATTATTTG TGGTGCCCAA CATTGCGGCT CTTGAGCCTG CTGCTGGTCT CCGGATGCCC 2040
AGTGAAGGTA TGTGGGATGG GGTGGTGGGG TAGGGGACGG TATCCTTTT TTGCTCCTAC 2100
50 TTGGAACAC CAAACACCCC AAGGAAGATG ATAGGCTCCA TCTTGGGCCA CCTGAGCTAT 2160
AGGCGAGCCT AATGGAATCA ACCATTCTG AGCACTAAAT GTATCATGAA AAGTGAATG 2220
GCCTGCTCAT AAGTTAGCT CATTCACTGG AAATGTAGAT TGATGTTCAA TGTAAACTG 2280
GAAGGAGCTT GGTGTGTGTG TCAGTGGTTA TATTAGTGGG TAGTGTAAAC TTTTATCCAG 2340
GTTGGGGTGA GGGGAGATGG CCACAGTAGC AAGTGGTGAC ACTAAATACC ATTTTGAAGG 2400
CTGATGTGTA TGTACATCAT TACTGTCCGT AGCAATGAAG GATACAGTAC TGTGTGTGG 2460
55 GTGAGTGTG CTATTGCCCA GCATTAAATAT TTGGTGTGTG ATGTTTGGAG CTATGAACA 2520
CGCAGGAGTG TTTTGTGCT ATTAATTTTA AGAGAAAGCA GCITTTTCTT AAAATCACT 2580
GTTGAGAAAC TTGCATGTCT GGAGGCGGTG TCCTCTCCGC CCGTGTGGGT CCGGATGAG 2640
TACGAGTTAT GGTCACTGCT ACAGCCTGAT CTCTTATGTG TTCAATAGCA TCGCTCTCC 2700
CATCAGAACT GTTGTGCTG AATGTGTTC TCTAGTTCTA GAAAATGACC ACTAATTTAA 2760
60 AAAAATCCGT TGTGAGTTT GCCCAGAGGC ACTGTGTCCA GAATTTCCCG TCCTGCTTCA 2820
GCCATGTCTT TGTCACTTGG CATTCTAAGC TAAAGCTTTA GCTTCCCAAT TCGTGTGTG 2880
CTAGGCCAAG ATTCGGGAGC TGTGCGCAGC CTGCTCAAAAT ATGGAAGAGA AACAACTGC 2940
GGTCAAAAGG GAGTGATTGG TTAAGTGGTG CGCGTCTATC TCATAACTAG ATGTACCAAC 3000
CAGGGAAGCG CCAAGGATGG AAAGGGGTAA CTTTGTGCT TCCAAAGTAG CTAAGCAGAA 3060
65 GTGGGGAGC AGTTTAGCCA GATGATCTTT GATTAGGCAA ACATTGAGTT TTAAGAGGC 3120
TGTCAGATTG AGGCCACTTG GTCCATTAGC TGGGGCAGCA AGATCACTAC TCAACGTTT 3180
CACACTGTGG CAAGATTGCT CTTCTAGTGG AATAATGCCC TAGTTTCTCT GAGATGATGT 3240
AAGTGGCATG ATGTTACCTA AGGCTTAGGC TTAGCTTGAT TTCTGGGCCC ACTGTCTGTG 3300
70 TTCTTAAGAT GCCAATCTGT TGCTTTTTTT TTTTCTTCC CCCATTAAA AGGATAGTAC 3360
CTACTCCCTC TAACCACCTC ACCCATTTCT TGAATGACAT TTTATCCTTC GGAAGAACAA 3420
AGGCTGTGAT GTAGTGACCTA TTGCTGTGCT CTCTGTGTG TGCTGTCTCT TGTCACAAA 3480
GTATTGCGG ACGTTGGATG CATTCAATTT CTGTAATAAA G

75 Seq ID NO: 306 Protein sequence:
Protein Accession #: NP_005333.1

1 11 21 31 41 51
80 MAKGDPPKKP GKMSAYAFV QTCREEHKKK NPEVPVNPAB FSKKCSERWK TMSGKEKSKP 60
DEMAKADKVR YDREMKDYGP AKGGKGGKDP NAPKRPSPGF FLFCSEFRPK IKSTNPGISI 120
GDVAKKLGEW WNNLNDSEKQ PYITKAARKL EKYEKDVADY KSKGKFDGAK GPAKVARKKV 180
EEEEEEEEEE EEEEEEEEE

85 Seq ID NO: 307 DNA sequence
Nucleic Acid Accession #: NM_022342
Coding sequence: 1..2178

1 11 21 31 41 51
5 ATGGGTACTA GGAAAAAGT TCATGCATT GTCCGTGTCA AACCCACGA TGACTTTGCT 60
CATGAAATGA TCAGATACGG AGATGACAAA AGAAGCATTT ATATTCACTT AAAAAGAGAC 120
ATTCCGAGAG GAGTTGTCAA TAACCAACAG ACAGACTGGT CGTTTAAGTT GGATGGAGTT 180
TTCAAGATG CCTCCAGGA CTGTGTTAT GAGACAGTTG CAAAGGATGT GGTTCCTCAG 240
CCCTCCGATG GCTATAATGG CACCATCATG TGTATGGGC AGACGGGAGC TGGCAAGACA 300
ACACCATGA TGGGGGCAAC TGAGAATTAC AAGCACCGGG GGATCCTCCC TGGTCCCTCG 360
10 AGCAGGTTT TTAGGATGAT CGAAGAACGC CCCACACATG CCATCACTGT GCGTGTTC 420
ACTTGGAAA TCTATAATGA GAGCCTGTTT GATCTCCTGT CCATCTGCCC CTATGTGGGA 480
CCTCAGTCA CACCAATGAC CATCGTGGAA AACCTCAAG: GAGTCTTCAT: TAAGGGCTTG 540
CAGTTCAACC TCACAAGTCA GGAGGAGGAT GCATTACGCC TCCTTTTGA: GGGTGAGACC 600
ACAGGATTA TAGCCTCCCA CACTATGAAC AAAAATCTTT CCAGATCACA: CTGCATTTTC 660
15 CCATCTACT TTAGAGGCCA TTCCCGGACC TTATCAGAGG AAAAGTACAT: CACTTCCAAA 720
TTAACTTGT TGGATCTGGC AGGCTCAGAG AGGCTGGGGA AGTCTGGTCT TGAGGGCCAA 780
TCCTGAAGG AAGCCACCTA CATCAACAAA TCGCTCTCAT: TCCTGGAGCA GGCCATCATT 840
CCCTTGGGG ACCAGAAGCG GGACCAATC CCCTTTCGGC AGTCAAGCT: CACCCAGGCT 900
20 TGAAGGACT CGTTAGGGGG AAATGCAAT: ATGGTCTCTG TGACAAACAT: CTATGGAGAA 960
CTGCCAGT TAGAAGAAAC GCTATCTTCA: CTGAGATTG: CCAGCAGGAT GAAGTAGTC 1020
CCACTGAGC TCGCCATCAA TGAAAAGTAT GATGCTGAGA: GAATGGTCAA: GAACCTGGAG 1080
AGGAACATG CACTACTCAA GCAGGAGCTG: GCTATCCATG ACAGCCTGAC CAACCGCACC 1140
TTGTGACCT ATGACCCCAT GGATGAAATC CAGATTGCTG AGATCAACTC CCAGGTGCGG 1200
25 GGTACCTGG AGGGGACACT GGACGAGATC GACATAATCA GCCTTAGACA: GATCAAGGAG 1260
TGTTCACCC AGTTCCGGGT GGTTCGAGC CAACAGGAAC AGGAAGTGA GTCCACTTTG 1320
GCAGGAAGT TGCATCTGAG TGACAGGAAT GACTTTGAGC CCATTTCTGC TATCCAGAAG 1380
CGGGGCTTG TGGATGTGA TGGCCACCTA GTGGGTGAGC CTGAAGGACA AAATTTTGA 1440
TCGGAGTGG CCCCTTTCTC TACCAAACTT GGAAGAAAAG CCAAGTCCAA: GAAGACATTC 1500
30 AAGAGCCAC TCAGGCCCGA CACCCACACC TCCAAACCG: TGGCCTTTGA: GGAGTTAAG 1560
ATGAGCAAG GTAGTGAGAT CAACCGAATT TTCAAGAAA ACAATCCAT: CTTGAATGAA 1620
GGAGGAAAT CGGCCAGCGA GACCCACACG CACATCAATG CCATCAAGCG GGAGATTGAT 1680
TGACCAAGG AGGCCCTGAA TTTCCAGAA TCACTACGGG AGAAGCAAGG CAAGTACGAA 1740
ACAAGGGGC TGATGATCAT CGATGAGGAA GAATTCCTGC TGATCCTCAA GCTCAAAGAC 1800
35 TCAAGAAGC AGTACCGCAG CGAGTACCAG GACCTGCGTG ACCTCAGGCG TGAGATCCAG 1860
ATTGCCAGC ACCTAGTGGG TCAGTGTGCG CACCGCTGCG TCATGGAATT TGACATCTGG 1920
40 ACAATGAGT CCTTTGTCT: CCTGAGGAC ATGCAGATGG: CACTGAAGCC AGGCGGCAGC 1980
TCCGCCAGC GCATGGTCCC TGTGAACAGG ATTGTGCTC TGGGAGAAGA TGACCAAGAC: 2040
AATTCAAGC AGCTGCAGCA GAGGTGCTT CCTGAGGGCC CTGATTCCAT: CTCCTTCTAC 2100
ATGCCAAG TCAAGATAGA GCAGAAGCAT AATTACTTGA AAACCATGAT GGGCTCCAG 2160
AGGCACATA GAAAAATG

Seq ID NO: 308 Protein sequence:
Protein Accession #: NP_071737

45 1 11 21 31 41 51
MGTRKKVHAP VRVKPTDDFA HEMIRYGGDK RSIDIHLKDD IRRGVVNNQQ TDWSFKLDGV 60
LHDASQDLVY ETVAKDVVSQ ALDGYNGTIM CYGQTGAGKT YTMGATENY: KHRGILPRAL 120
50 QQVFRMIEER PTHAITVRVS YLEIYNESLF DLLSTLPYVG PSVTMTIVE NPQGVFIKGL 180
SVHLTSQEDD AFSLFFEGET NRITASHTMN: KNSRSCHIF TIYLEAHSRT LSEKYITSK 240
INLVDLAGE RLKSGSGEG: VLKEATYINK: SLSPLEQAI: ALGDQKRDI PFRQCKLTHA 300
LKDSLGNMNC MVLVTNIYGE AAQLEETLSS LRFASRMKLV: TTEPAINEKY DAERMVKNLE 360
KELALLKQEL AIHDSLNRTE FVTYDPMDEI: QIABINSQVR RYLESTLDEI: DIISLRQIKE 420
VFNQFRVVL SQQEQEVESLT RRYKTLIDRN: DFAAISAIQK: AGLVDVDGHL VGEPEGQNF 480
55 LGVAPFSTKP GKAKSKKTP KEPLRPDTPP: SKPVAFEFFK: NEQGSSEINRI FKENSILNE 540
RRKRASSTQ HINAIKREID VTKBALNFQK SLREKQKYE: NKGLMIDEE EPLILKLD 600
LKQYRSEVQ DLRLRAEIQ YCOHLVDQCR HRLMEFDIY YNESFVIPED MQMALKPQGS 660
IRPMGVVVRN IVSLGEDDQD KFSQLQORVL PEGPDSISFY NAKVKIEQKH NYLTKMGLQ 720
QAHKK

Seq ID NO: 309 DNA sequence
Nucleic Acid Accession #: CAT cluster

65 1 11 21 31 41 51
TTTTTTTTT TTTTTTTAA TGCCTGCTGT CATGCTCTGT CTACCAGGCT GAATTTCCAA 60
AAATTTCTGC ATAGCAATTT TAGCCAAAAC TATATATGTT CTGGGGAGGA TAGGCATAGG 120
CACATTGAAG ACCAAAGGAA AGAGTGAAGA AGTGTAGTTG GGTCAATGTG AATGGATGTT 180
70 TAGATTGTCA AGAAAAGTGG GCCAGAGGCC CCACCTCACA CTAGGACGGC AATGCCTCT 240
CAATTAGTATC TCAGGCACCA TGGGTCTTAT TTGGTGTCT: AAGAAACACC CTCACAAAAG 300
TAATGAACCC TCAGCCTCCA GCTTCTCTTC TTCGGGATTG TTCTTAGGGC CTCCTTTTTC 360
CTTTTATGTT TCCAGTACCC TGAATTTCTT: ATTCCCATCC CCATTAAAA TCTGCTTCAA 420
AGAAAAACA AGAAGGACAC ATTCACCTTA AGATCCAAAT GAATGATAAG AGCTTAAAC 480
75 ATTATACTTA TCAGTATTAT TTGCATTTTT ATAGAAACCA AAACCATATT TCAACAC 540

Seq ID NO: 310 DNA sequence
Nucleic Acid Accession #: NM_018622.2
Coding sequence: 1-1140

80 1 11 21 31 41 51
ATGGCGTGGC GAGGCTGGGC GCAGAGAGGC TGGGGCTGCG GCCAGGCGTG GGGTGGTGG 60
GTGGGGGSCC GCAGCTGCGA GGAGCTCACT CGGCTCCTAA CCCGCGGCA GCTCCTCGGA 120
85 CGCAGGTTTA ACTTCTTTAT TCAACAAAAA TGGCGATTCA GAAAGACACC CAGGAAGGTT 180
GAACCTCGAA GATCAGACCC AGGGACAAGT GGTGAAGCAT ACAAGAGAAG TGCTTTGATT 240
CCTCCTGCG AAGAAACAGT CTTTATCCT TCTCCTATC CTATAAGGAG TCTCATAAAA 300
CCTTTATTTT TTACTGTGG GTTTACAGGC TGTGATTGG GATCAGCTGC TATTGGCAA 360

5
10
15

TATGAATCAC TGAATCCAG GGTCCAGAGT TATTTTGATG GTATAAAAGC TGAATGGTTG 420
GATAGCATAA GACCACAAA AGAAGGAGAC TTCAGAAAAG AGATTAAACAA GTGGTGGAAT 480
AACCTAAGTG ATGGCCAGCG GACTGTGACA GGTATTATAG CTGCAATGTG CTTGTATTTC 540
TGTTTATGGA GAGTACCTTC TCTGCAGCGG ACAAATGATCA GATATTTTAC ATCGAATCCA 600
GCCTCAAAGG TCCTTTGTTC TCCAAATGTTG CTGTCAACAT TCAGTCACTT CTCTTATTT 660
CACATGGCAG CAAATATGTA TGTTTTGTGG AGCTTCTCTT CCAGCATAGT GAACATTCTG 720
GGTCAAGAGC AGTTCATGGC AGTGTACCTA TCTGCAGGTG TTATTTCCAA TTTTGTCACT 780
TACCTGGGTA AAGTTGCCAC AGGAAGATAT GGACCATCAC TTGGTGATC TGGTGCCATC 840
ATGACAGTCC TCGCAGCTGT CTGCACTAAG ATCCCAAGAG GGAGGCTTGC CATTATTTTC 900
CTTCCGATGT TCACGTTTAC AGCAGGGAAT GCCCTGAAAG CCATTATCGC CATGGATACA 960
GCAGGAATGA TCCTGGGATG GAAATTTTIT GATCATGCGG CACATCTTGG GGGAGCTCTT 1020
TTTGGAAATAT GGTATGTATC TTACGGTCAT GAACGTATTT GGAAGAACAG GGAGCCGCTA 1080
GTGAAATCT GGCATGAAAT AAGGACTAAT GGCCCCAAAA AAGGAGGTGG CTCTAAGTAA

Seq ID NO: 311 Protein sequence:
Protein Accession #: NP_061092.2

20
25
30

1 11 21 31 41 51
MAWRGWAQRG WGGCQAWGAS VGGRSCEELT AVLTPPQLLG RRFNFFIQK CGFRKAPRKV 60
EPRRSDFGTS GEAYKRSALI PPVEETVFYP SPYPIRSLIK PLFPTVGFTG CAFGSAAIWQ 120
YESLSKRVQS YFDGIKADWL DSIRPQKEGD FRKEINKWNW NLSGQRTVT GIIAANVLVF 180
CLWRVPSLQR TMIRYFTSNP ASKVLCSFML LSTFHSFSLF HMAANMYVLW SFSSSIUNIL 240
QEQFMAVYL SAGVINFPVS YLGKVTATRY GPSLGAAGAI MTVLAAVCTK IPEGLRIIIP 300
LPMFTPTAGN ALKAIAMDT AGMILGWKPF DHAHLGGAL PGIVYVYGH ELIWNKREPL 360
VKIWHIRTN GPKKGGGSK

Seq ID NO: 312 DNA sequence
Nucleic Acid Accession #: NM_000625
Coding sequence: 195..3656

35
40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
CTCTCGGCCA CCTTTGATGA GGGGACTGGG CAGTTCTAGA CAGTCCCGAA GTTCTCAAGG 60
CACAGGTCTC TTCTCGGTTT GACTGTCTTT ACCCCGGGGA GGCACTGCAG CCAGCTGCAA 120
GCCCCACAGT GAAGAACATC TGAGCTCAAA TCCAGATAAG TGACATAAGT GACCTGCTTT 180
GTAAAGCCAT AGAGATGGCC TGTCCTTGGA AATTTCTGTT CAAGACCAAA TTCCACCACT 240
ATGCAATGAA TGGGGAAAAA GGCATCAACA ACAAATGTGA GAAAGCCCCC TGTGCCACCT 300
CCAGTCCAGT GACACAGGAT GACCTTCAGT ATCAACAACCT CAGCAAGCAG CAGAATGAGT 360
CCCCGAGGCC CCTCGTGGAG ACGGGAAAGA AGTCTCCAGA ATCTCTGGTC AAGCTGGATG 420
CAACCCCATC GTCTCTCCCA CGGCATGTGA GGATCAAAA CTGGGGCAGC GGGATGACTT 480
TCCAAGACAC ACTTCACCAT AAGGCCAAAG GGATTTTAAC TTGCAGGTCC AAATCTTGCC 540
TGGGGTCCAT TATGACTCCC AAAAGTTTGA CCAGAGGACC CAGGGACAAG CCTACCCCTC 600
CAGATGAGCT TCTACCTCAA GCTATCGAAT TTGTCAACCA ATATTACGGC TCCCTCAAAG 660
AGGCAAAATC CAGAGAAATC CTGGCCAGGG TGGAGGCGGT AACAAAGGAG ATAGAAACAA 720
CAGTAACCTA CCAACTGAGC GGAGATGAGC TCATCTTCGC CACCAAGCAG GCCTGGCGCA 780
ATGCCCCAGC CTGCATTGGG AGGATCCAGT GGTCCAACCT GCAGGTCTTC GATGCCCGCA 840
GCTGTTCCAC TGCCCGGGAA ATGTTTGAAC ACATCTGCAG ACACGTGCGT TACTCCACCA 900
ACAAATGGCA CATCAGGTGC GCCATCACCG TGTTCCCCCA GGGGAGTGAT GGCAAGCACG 960
ACTTCCGGGT GTGGAAATGCT CAGCTCATCC GCTATGTGCG CTACCCAGATG CCAGATGGCA 1020
GCATCAGAGG GGAACCTGGC AACGTGGAAT TCACTCAGCT GTGCATCGAG CTGGGCTGGA 1080
AGCCCAAGTA CGGCCGCTTC GATGTGCTCC CCCTGCTCCT GCAGGCCAAT GGCCGTGACC 1140
CTGAGCTCTT CGAAATCCCA CCTGACCTTG TGCTTGAGGT GGCCATGGAA CATCCCAAAT 1200
ACGAGTGGTT TCGGGAACTG GAGCTAAAGT GGTACGCCCT GCGTGCAGTG GCCAACATGC 1260
TGCTTGAGGT GGGCGGCTG GAGTTCCAGG GGTGCCCTTC CAATGGCTGG TACATGGGCA 1320
CAGAGATCGG AGTCCGGAGC TTCTGTGATG TCCAGCGCTA CAACATCCTG GAGGAAGTGG 1380
GCAGGAGAAAT GGGCGCTGAA ACGCACAAAGC TGGCTCGCT CTGGAAAGAC CAGGCTGTGG 1440
TTGAGATCAA CATTGTCTGT TCCATAGTT TCCAGAAGCA GAATGTGACC ATCATGGACC 1500
ACCACTCGGC TGACAGATCC TTCATGAAGT ACATGCAGAA TGAATACCGG TCCCGTGGGG 1560
GCTGCCCGGC AGACTGGATT TGGCTGTGCT CTCCTCATGC TGGGAGCATC ACCCCGCTGT 1620
TTCACCAAGG GATGCTGAAC TACGTCTGT CCCCTTCTA CTACTATCAG GTAGAGGCCT 1680
GGAAAACCCA TGTCTGGCAG GACGAGAAGC GGAGACCCAA GAGAAGAGAG ATTCCATTGA 1740
AAGTCTTGGT CAAAGCTGTG CTCTTTGCC TGTATGCTGAT GCGCAAGACA ATGGCGTCCC 1800
GAGTCAGAGT CACCATCCTC TTTGCGACAG AGACAGGAAA ATCAGAGGCG CTGGCCTGGG 1860
ACCTGGGGGC CTTATTGAGC TGTGCTTCA ACCCCAAAGT TGTCTGCATG GATAAGTACA 1920
GGCTGAGCTG CTTGGAGGAG GAAAGGCTGC TGTGTGGTGT GACCAGTACG TTTGGCAATG 1980
GAGACTGCCC CCGCAATGGA GAGAACTGA AGAAATCGCT CTTCATGCTG AAAGAGCTCA 2040
ACAACAAAT CAGGTACGCT GTGTTTGGCC TCGGCTCCAG CATGTACCCT CGGTTCTGG 2100
CCTTTGCTCA TGACATTGAT CAGAAGCTGT CCCACCTGGG GGCTCTCAG CTCACCCCGA 2160
TGGGAGAAGG GGATGAGCTC AGTGGGCAGG AGGACGCTT CCGCAGCTGG GCGGTGCAAA 2220
CCTTCAAGGC AGCTGTGAG ACGTTTGATG TCCAGGCAAA ACAGCACATT CAGATCCCCA 2280
AGCTCTACAC CTCCAATGTG ACCTGGGACC CGCACCACTA CAGGCTCGTG CAGGACTCAC 2340
AGCCTTTGGA CCTCAGCAA GCCCTCAGCA GCATGCATGC CAAGAAGCTG TTCACCATGA 2400
GGCTCAAATC TCGGCAAGAT CTACAAAGTC CGCATCCAG CCGTGCCACC ATCTCTGTGG 2460
AACTCTCTG TGAGGATGGC CAAGGCTGTA ACTACTGCC GGGGGAGCAC CTGGGGTTT 2520
GCCCGAGCAA CCAGCCGGCC CTGGTCCAAG GCATCTGGA GCGAGTGGTG GATGGCCCCA 2580
CACCCACCA GGCAGTGCGC CTGGAGGCC TGGATGAGAG TGGCAGCTAC TGGGTCACTG 2640
ACAAGAGGCT GCCCCCTGTC TCACTCAGCC AGGCCCTCAC CTACTTCTG GACATCACA 2700
CACCCCAAC CACGCTGCTG CTCCAAAAGC TGGCCAGGT GGCCACAGAA GAGCCTGAGA 2760
GACAGAGGCT GGAGSCCTG TGCCAGCCCT CAGAGTACAG CAAGTGGAGG TTCACCAACA 2820
GCCCCACATT CCTGGAGGTG CTAGAGGAGT TCCGTCCTCT GCGGTGTCT GCTGGCTTCC 2880
TGCTTTCCCA GCTCCCATTT CTGAAGCCCA GTTCTACTC CATCAGCTCC CCCCAGGATC 2940
ACACGCCCAT GGAGATCCAC CTGACTGTGG CCGTGGTCAC CTACCACACC CGAGATGGCC 3000
AGGTCCCT GCACACGGC GTCTGCAGCA CATGGCTCAA CAGCCTGAAG CCCCAGGACC 3060
CAGTCCCTG CTTTGTGCGG AATGCCAGCG GCTTCCACCT CCCCAGGAT CCCCCTCATC 3120

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85

```

CTTCATCTCT CATCGGCGCT GGCACAGSCA TOGCGCCCTT COGCAGTTTC TGGCAGCAAC 3180
GGCTCCATGA CTCCAGCAC AAGGGAGTGC GGGGAGGCGG CATGACCTTG GTGTTTGGGT 3240
GCGCGCGGCC AGATGAGGAC CACATCTACC AGGAGGAGAT GCTGGAGATG GCCCAGAAGG 3300
GGGTGCTGCA TCGCGTGCAC ACAGCCTATT CCGCGCTGCC TGGCAAGCCC AAGGTCTATG 3360
TTCAGGACAT CCTCGGCGAG CAGCTGGCCA GCGAGGTGCT COGTGTGCTC CACAAGGAGC 3420
CAGGCCACCT CTATGTTTGC GGGGATGTGC GCATGGCCCG GGAGGTGGCC CACACCTGA 3480
AGCAGCTGGT GGCTGCCAAG CTGAAATTGA ATGAGGAGCA GGTGGAGGAC TATTCTTTTC 3540
AGCTCAAGAG CCAGAAGCGC TATCAGGAAG ATATCTTTGG TGCTGTATTT CCTTACGAGG 3600
CGAAGAAGGA CAGGGTGGCG GTGCAGCCCA GCAGCCTGGA GATGTCAGCG CTCTGAGGGC 3660
CTACAGGAGG GGTAAAGCT GCCGGCACAG AACTTAAGGA TGGAGCCAGC TCTGCATTAT 3720
CTGAGGTGAC AGGGCCTGGG GAGATGGAGG AAGTGTATAT CCCCCAGCCT CAAGTCTTAT 3780
TTCCTCAAGG TTGCTCCCA TCAAGCCCTT TACTTGACCT CCTAACAAAGT AGCACCCCTG 3840
ATTGATCGGA GCCTC
  
```

Seq ID NO: 313 Protein sequence:
 Protein Accession #: NP_000616

1 11 21 31 41 51
 20
 25
 30
 35
 40

```

MACPWKFLPK TKFHQYAMNG EKGINNVEK APCATSSPVT QDDLYHNLS KQNESQPQL 60
VBTGKKSPES LVKLDATPLS SPRHVRKKNW GSGMTFQDTL HHKAGILTC RSKSCLGSIM 120
TPKSLTRGPR DKPTPPDELL PQAIEFVNQY YGSLKEAKIE EHLARVEAVT KEIETTVTYQ 180
UTGDELIPAT KQAWNRNAPRC IGRIQWSNLQ VPDARSCSTA REMFEHICRH VRYSTNNNGNI 240
RSAITVPPQR SDGKHDFRVW NAQLIRYAGY QMPDGSIRGD PANVEFTQLC IDLGWPKPYG 300
RFDVVVLVLQ ANKRDPELFE IPDLVLLEVA MEHPKYEWFR EELKLYALP AVANMLLEVQ 360
GLEFPFGCPN GWYMTGIGV RDFCDVQRYN LEEVGRRMG LETHKLASLW KDQAVVEINI 420
AVLHSPQKQN VTIMDHSAE ESPMKYQNE YRSRGCCPAD NIWLVPMSG SITPVFHQEM 480
LWYVLSPPFY YQVEAMKTHV WQDEKRRPKR REIPLKVLVK AVLFACLMLR KTMASRVRT 540
ILFATETGKS EALAWDLGAL FSCAPNPKVV CMQKRLSCL EERLLLVVT STFGNGDCPG 600
NGEKLKSLF MLKELNNKFR YAVFGLGSSM YPRFCAPAH IDQLSHLGA SQLTPMGEGD 660
ELSGQEDAFR SWAYQTFKAA CETFDVRGKQ HIQIPKLYTS NVTWDPHYR LVQDSQPLDL 720
SKALSSMHAK NVFTMLRKR QNLQSPSSR ATILVELSCE DGQGLNLYLP EHLGVCPGNQ 780
PALVQGILER VVDGTFPHQA VRLEALDESG SYWVSDKRLP PCSLSQALTY FLDITTPPTQ 840
LLLQKLAQVA TEPPERQRL EALQPSSEYSK WKFTNSPTFL EVLEEFPSLR VSAGFLLSQL 900
PILKPRYSI SSPRDHTPTE IHLTVAVVTY HTRDQGGPLH HGVCSTWLSN LKPDQVPVPCF 960
VRNAGSGPHL EDPGSHPCIL GPGTGIAPFR SFWQORLHDS QHKGVRGGRM TLVFGCRRPD 1020
BDHIYQSEML EMAGKGVLHA VHTAYSRLPG KPKVYVQDIL RQQLASEVLR VLHKEPGHLY 1080
VCGDVRMARD VAHLKQLVA AKLKLNEEQV EDYFPQLKSQ KRYHEDIFGA VPPYEAKKDR 1140
VAVQPSLEM SAL
  
```

Seq ID NO: 314 DNA sequence
 Nucleic Acid Accession #: XM_087254
 Coding sequence: 47..2332

45
 50
 55
 60
 65
 70
 75
 80
 85

```

1 11 21 31 41 51
AGAGTACGTC TTTACAGATA AAATCGGTAC ACTGACAGAA AATGAGATGC AGTTTCGGGA 60
ATGTTCAATT AATGGCATGA AATACCAAGA AATTAATGGT AGACTTGTAC CGSAAAGGACC 120
AACACCAAGC TCCTCAGAAG GAAACTTATC TTATCTTAGT AGTTTATCCC ATCTTAACAA 180
CTTATCCCAT CTACAAACCA GTTCTCTTTT CAGAACCAGT CCTGAAAATG AAACCTGAAC 240
AATTAAAGAA CATGATCTCT TCTTTAAAGC AGTCAGTCTC TGTCACACTG TACAGATTAG 300
CAATGTTCAA ACTGACTGCA CTGGTGATGG TCCCTGGCAA TCCAACCTGG CACCATGCA 360
GTTGGAGTAC TATGCATCTT CACCAGATGA AAAGGCTCTA GTAGAAGCTG CTGCAAGGAT 420
TGGTATTGTG TTTATTGGCA ATTCTGAAGA AACTATGGAG GTTAAACTC TTGGAAGACT 480
GGAAACGGTAC AAACCTGCTC ATATTCTGGA ATTTGATTCA GATCGTAGGA GAATGAGTGT 540
AATTGTTTCA GCACCTTCAG GTGAGAAAGT ATTTATGCTT AAAGGAGCTG AGTCATCAAT 600
TCTCCCTAAA TGTATAGGTG GAGAAATAGA AAAAACCCAGA ATTCATGTAG ATGAATTGTC 660
TTTGAAAGGG CTAAGAACTC TGTGTATAGC ATATAGAAAA TTTACATCAA AAGAGTATGA 720
GGAAATAGAT AAACGTCATC TTGAAGCCAG GACTGCCTTG CAGCAGCGGG AAGAGAAATT 780
GGCAGCTGTT TCCAGTTTCA TAGAGAAAGA CCTGATATTA CTGGAGCCA CAGCAGTAGA 840
AGACAGACTA CAAGATAAAG TTGAGAAAC TATTGAAGCA TTGAGAAAGG CTGGTATCAA 900
AGTATGGGTA CTACTGGGG ATAAACATGA AACAGCTGTT AGTGTGAGTT TATCATGTGG 960
CCATTTTCAT AGAACCATGA ACATCCTTGA ACTTATAAAC CAGAAATCAG ACAGCGAGTG 1020
TGCTGAACAA TTGAGGACGC TTGCCAGAAG AATTACAGAG GATCATGTGA TTCAGCATGG 1080
GCTGGTAGTG GATGGGACCA GCCTATCTCT TGCACTCAGG GAGCATGAAA AACTATTAT 1140
GGAAAGTTTG AGAAATGTGT CAGCTGTATT ATGCTGTCTG ATGGCTCCAC TGCAGAAAGC 1200
AAAAGTAATA AGACTAATAA AAATATCACC TGAGAAACCT ATAACATTGG CTGTTGGTGA 1260
TGGTGCTAAT GACGTAAGCA TGATACAAGA AGCCCATGTT GGCATAGGAA TCATGGGTAA 1320
AGAAAGGAAG CAGGTCGCAA GAAACAGTGA CTATGCAATA GCCAGATTTA AGTTCTCTCT 1380
CAAAATTGCT TTTGTTTCATG GTCATTTTGA TTATATTAGA ATAGCTACCC TTGTACAGTA 1440
TTTTTTTAT AAGAAATGTG GCTTTATCAC ACCCCAGTTT TTATATCAGT TCTACTGTTT 1500
GTTTTCTCAG CAACATTGT ATGACAGCGT GTACCTGACT TTATACAATA TTTGTTTAC 1560
TTCCCTACCT ATTCTGATAT ATAGTCTTTT GGAACAGCAT GTAGACCTTC ATGRTTTACA 1620
AAATAAGCCC ACCCTTTATG GAGACATTAG TAAAAACCGC CTCTTAAGTA TTAATAACAT 1680
TCTTTATTGG ACCATCTCGG GCTTCAGTCA TGCCTTTATT TTCTTTTGG GATCCTATT 1740
ACTAATAGGG AAGATACAT CTCTGCTTGG AAATGGCCAG ATGTTTGGAA ACTGGACATT 1800
TGGCATTGTT GTCTTCACAG TCATGGTTAT TACAGTCACA GTAAAGATGG CTCTGGAAAC 1860
TCATTTTGG ACTTGGATCA ACCATCTCGT TACCTGGGGA TCTATTATAT TTTATTTTGT 1920
ATTTTCTCTG TTTATTGGAG GGATTCTCTG GCCATTTTGG GGCTCCAGCA ATATGTATT 1980
TGTGTTTAT CAGCTCTCTG CAAGTGGTTC TGCTTGGTTT GCCATAATCC TCATGGTTGT 2040
TACATGTCTA TTTCTTGATA TCATAAAGAA GGTCTTTGAC CGACACCTCC ACCCTACAAG 2100
TACTGAAAAG GCACAGCTTA CTGAAACAAA TGCAGGTATC AAGTGTCTGG ACTCCATGTG 2160
CTGTTTCCCG GAAGGAGAAG CAGCGTGTGC ATCTGTGGGA AGAATGCTGG AACGAGTTAT 2220
AGGAAGATGT AGTCCACCC ACATCAGCAG ATCATGGAGT GCATCGGATC CTTTCTATAC 2280
CAACGACAG AGCATCTTGA CTCTCTCCAC AATGGACTCA TCTACTTGT AAAGGGGCG 2340
  
```

5
10
15
20
25
30
35
40
45
50
55
60

TAGTACTTTG TGGGAGCCAG TTCACCTCCT TTCCTAAAT TCAGTGTGAT CACCCGTGTA 2400
ATGGCCACAC TAGCTCTGAA ATTAATTTCC AAAATCTTTG TAGTAGTTCA TACCCACTCA 2460
GAGTTATAAT GGCAAAACAA CAGAAAGCAT TAGTACAAGC CCCTCCCAAC ACCCTTAATT 2520
TGAATCTGAA CATGTTAAAA TTGAGAGATA AAGAGACATT TTTCATCTCT TTGTCTGGTT 2580
TGTCCTTGT GCTTATGGGA CTCCTAATGG CATTTTCAGT CTGTGCTGAG GCCATTATAT 2640
TTTAATATAA ATGTAGAAAA AAGAGAGAAA TCCTAGTAAA GAGTATTTT TAGTATTAGC 2700
TTGATTATG ACTCTTCTAT TTAATCTGCT TTCTGTAATAT TATGCTGAAA GTTTGCTCTG 2760
AGAACTCTAT TTTTATTATA GAGTTATATT TAAAGCTTTT CATGGGAAAA GTTAATGTGA 2820
TACTGAGGA ATTTTGTGCC CTCAGTGACC TGTGTTGTTA ATTCATTAAT GCATTCTGAG 2880
ATACAGAGC AAATTAGGAG AATCATTTCC AACCATTAAT TACTGCAGTA TGGGGAGTAA 2940
ATTTATACCA ATTCCTCTAA CTGTACTGTA ACACAGCCTG TAAAGTTAGC CATATAAATG 3000
CAAGGGTATA TCATATATAC AAATCAGGAA TCAGGTCCTG TCACCGAATC TCAAAATTGAT 3060
GTTTACTAAT ATTTTGTGTA CAGAGTATAA AGACCCATA GTGGGTAAAT TAGATACTAT 3120
TAGCATATTA TTAATTTAAT GTCTTATCA TTGGATCTTT TGCGATGCTT AATCTGGTTA 3180
ACATATTTAA ATTTGCTTTT TTTCTCTTA CCTGAAGGCT CTGTGTATAG TATTTATGTA 3240
CATGTTGTA CAGTTTAACT ATATCAATAA AAGTTTGGGA CAGTATTAA ATATTGCAA 3300
TATGTTTAA TATACAAATC AGAATAGTAT GGGTAATTA ATGAATACAA AAGAAGAGC 3360
CTCTTCTGC AGCGGACTTA GACATGCTCT TCCCTTTCTA TAAGCTAGAT TTTAGATAA 3420
AGGGTTTCAG TTAATTAATCT TATTTTCAGG TTATGTCATC TAACTTATAG CAAACTACCA 3480
CAATACAGTG AGTTCTGACA GTGTCCAGT ACAAGGCATA TTTCAGGTGT GGCTGTGGA 3540
TGTAATAATG CTCGAATCTG ATCAGGTAAT GTTAGCAATA AATTAATGTC TAAGAAATGAT 3600
TAATCGGGTA CATGTTACTG TAAATTAATC ATTCGACTTC AAAACCTAAC TTCCATCTG 3660
AATTATCAA GTAGTTTCAAT ATTGTCTATT GTTTTGTGTT TATTGAAAAG TAATGTTGTC 3720
TTAAGATTTA GAAGTGATTA TTAGCTTGAG AACTATTACC CAGCTCTAAG CAAATAATGA 3780
TTGTATACAT ATTAAGATAA TGGTTAAATG CGGTTTACC AAGTTTCCC TTGAAAATGT 3840
AATTCCTTTA TGGAGATTTA TTGTGCAACC CTAAGCTTCC TTCCCATTTT ATGAATATAA 3900
GGCTTCTAGA ATTCGACTCG CAGGGGAAAG AATGGTAGAG ACAGAAATTA AGACTTTATC 3960
CTGTGTTGCT TGTAACTAT TATTTCTTG CTAATGTAAC ATTTGCTGCT TCCAGTGATG 4020
TAAGGATATT AAGTTATTA GCTAAATATT AATTTTCAAA AATAGTCCCT CTTTAACTTA 4080
GATATTTTAT AGCTGGATT AGGAAGATCT GTTATCTGG AAGTACTAAA AAGAATAATA 4140
CAAGGTACAA TGTCTGCATT CACTAATTC TGTCCAGAA GAGGAAATAA TGAAGATATA 4200
CTCAGTAGAG TACTAGTGGG GAGGATATGG AAAATTGCTC ATAAAACTCT TTATAAAACG 4260
TGCAATAAC AAAATGACAC CCAGTAGGCC TGCATTACAT TTACATGACC GTGTTTATTT 4320
GCCATCAAT AAATCTAGTA CTGACACCAG ACAAGACTC CAAAGTCATA AAATAGCCTA 4380
TGACCAACTG CAGCAAGACA GGAGGTGAGC TCGCTATAA TGGTGCTTAA AGTGTGATTG 4440
ATGTAATTTT CTGTACTCAC CATTGGAAGT TAGTTAAGGA GAACCTTATT TTTTAAAAA 4500
AAGTAAATGG CAACCACTAG TGTGCTCATC CTGAACCTGT ACTCCAAATC CACTCCGTTT 4560
TTAAAGCAAA ATTATCTGT GATTTTAAAG AAAGAGTTT CTATTTATT TAAAGAGTAA 4620
CAATGCAATG TGCAAGCTTT CAGTAGTTT CTAGTGCTAT ATTATCCTG TAAACTCTT 4680
ACTACGTAAC CAGTAATCAC AAGGAAAGTG TCCCTTTGCT ATATTTCTT AAAATCTTT 4740
CTTTGGAAGG TATGATGTTG ATAATTAATC TACCCTTATC TGCCAAACC AGAGCAAAAT 4800
GCTAAATACG TTTATGCTAA TCAGTGGTCT CAAATCGATT TGCCCTCCTT TGCCCTCTCA 4860
GAGGGCTGTA AGCCTGAAGA TAGTGGCAAG CACCAAGTCA GTTTCCAAAA TGCCCCCTCA 4920
GCTGCTTAA GTGACTCAGC ACCCTGCTC AGCTTCAGCA GGCGTAGGCT CACCTGGGGC 4980
GGAGCAAAT ATGCGCCAGG GAGAACTACA GCTACGAAGA CCTGCTGTCG AGTTGAGAAA 5040
AGGGGAGAAT TTATGGTCTG AATTTTCTAA CTGCTCTCT TCTTGGGTCT AAAGCTCATA 5100
ATACACAAAG GCTTCCAGAC CTGAGCCACA CCCAGGCCCT ATCCCTGAACA GGAGACTAAA 5160
CAGAGGCAAA TCAACCTAG GAAATACTTG CATCTGCCC TACGGTTAGT ACCAGGACTG 5220
AGGTCAATTC TACTGGAACA GATTGTGAGA TTGAACCTAT CTGATCGCTT GAGACTCCTA 5280
ATAGGCAGGA GTCAAGGCAA CTAGAAATG GACAGTTAAG AGCCAAAGT TTTTAAATA 5340
TGCTACTCTG AAAAATCTCG TGAAGGCTGT AGGAAAAGGG AGAATCTTCC ATGTTGGTGT 5400
TTTTCCTGTA AAGATCAGTT TGGGTATGA TATAAGCAGG TATTAATAAA AATAACACAC 5460
CAAAGAGTTA CGTAAACAT GTTTTATTAA TTTTGGTCCC CACGTACAGA CATTTTATTT 5520
CTATTTGTA ATGAGTTAT TATTTTATAA AAGTAAAC ACTATTAAG TGCTGTTTAA 5580
TGTGAAATAA CTGAATGTT GTTCTATAA AAAATAGATC ATAACCTCAT ATATGTTTGT 5640
AATCATGGTA ATTTAGATTT TTATGAGGAA TGAGTATCTG GAAATATGT AGCAATACTT 5700
GGTTTAAAT TTTGACCTG AGACACTGTG GCTGTCTAAT GTAATCCTTT AAAAATCTC 5760
TGCAATGTA GTAAATGTAG TATATTATG TACAGCTACT CATAATTTT TAAAGTTTAT 5820
GAAGTTATAT TTAATCAATA AAAACTTTCC TATAT

Seq ID NO: 315 Protein sequence:
Protein Accession #: XP_087254

65
70
75
80

1 11 21 31 41 51
MQFRECSING MKYQEIINGRL VPEGPTPDSS EGNLSYLSL SHLNNLSHLT TSSSFRTSPE 60
NETELIKEHD LFFKAVSLCH TVQISNVQTD CTGDGPWQSN LAPSQLEYA SSPDEKALVE 120
AARIGIVFI GNSSEETMEVK TLGKLERYKL LHILEFDSR RMSVIVQAP SGEKLLFAKG 180
AESSILPKCI GGEIEKTRIH VDEFALKGLR TLICIAIRKFT SKEYEIDKR IFEARTALQO 240
REEKLAAPVQ FIEKDLILLG ATAVEDRLQD KVRETIENLR MAGIKVWVLT GDKHETAVSV 300
SLSCGHPHRT MNIELINQK SDSECAEQLR QLARRITEDH VIQHGLVVDG TSLSLALREH 360
EKLFMVCRN CSAVLCCRMA PLQKAKVIRL IKISPEKPIT LAVGDGANDV SMIEQAHVGI 420
GIMGKEGRQA ARNSDYALIR FKFLSKLLPV HGHFYIRIA TLVQYFFYKY VCFITPQPLY 480
QFYCLFSQQT LKDSVYLTLF NICFTSLPIL IYSLLEQHDV PHVLQNKPTL YRDISKNRLL 540
SIKTFLYNTI LGFSHAFIFP FGSYLLIGKD TSLLGNGQMF GNWTFGLTVP TVMVTITVTK 600
MALEHFWNTW INHLVTWGS IYFVPSLFY GGILWPFLGS QNMYVPFIQL LSSGSANFAI 660
ILMVVTCFLF DIIRKVFDRH LHPTSTERAQ LTETNAGIKC LDMSCCFPEG EAACASVGRM 720
LERVIGRCSP THISRSWSAS DPFTYNDRSI LTLSTMDSS C

Seq ID NO: 316 DNA sequence
Nucleic Acid Accession #: NM_004473
Coding sequence: 661..1791

85

1 11 21 31 41 51
CTCGCCAGCG GTCCCGGGG CTGGAGACCC ACGCCGTGGA GAGGACCAGC CTCAGGTCGC 60

CCGGCTGGG CCGGCGCCCC GACCTCGCTG CCCCAGCCTC GCGTCTCTGC CCGTGGCGCT 120
 TACCGCCACC TTGGCCCTCGG GGGCAGGGCA TGGGCGGGCC CCGCAGATC GCCCAGCGCC 180
 AGTACTAACT GCCCTGCTC TGGCCTTCGA GCCCGAAGCC TCTTCTGCGC GCACAACTTA 240
 5 GGCAGTAATC CTAACTAGC GGGCACCACA GACCAGTGC AGCCACCCCA ACCCAGGGAT 300
 CACTTCCGGA CCCCCTGACC GCCCGGCACC AGCGCGCAAG GGACCTTCA CCGCGAGACC 360
 AGAGTCCAGT CCGGCTGCGG AGGCCACCGC CGCTGCCCGC CTCGAGAAGC ACAACGCGGG 420
 CTGAGCGCTC GGCCTAGCGG TCACTCCCGA GCGCTGTCTT GCACGCGGCC AGCCCCAGAC 480
 CACGACGCTC GAGCTCTCAG CGCGGCGCAG CCGTGGCGGC TGGGCTCTCC GGGCCAGCCC 540
 10 GCGACGATCC CCTGAGCTCT CCGCAGAAGG GCGAGCGTTC CGTTCCGGGG ACGCCAGGCC 600
 CGCCCCCGCC CGCCGACAGC CGCGGGGATC CAGAGCCCGG GGGTGGCGGA CGCCCCCGCC 660
 ATGACTGCGC AGAGCGGGCC GCCGCGCGCG CAGCGCGAGG TGCTGGCTAC CGTGAAGGAA 720
 GAGCGCGCGC AGACGCGAGC AGGGGCGCGG GTCCAGGGG AGGCCACGGG CCGCGGGGCG 780
 GGGCGGGGCG GCGCAGAGCG CCCCCTGCAG CGCGGGAAGC CGCCTACAG CTACATCGCG 840
 15 CTCATCGCCA TGGCCATCGC GCACGCGCCC GAGCGCGGCC TCAAGCTGGG CGGCATCTAC 900
 AAGTTTATCA CCGAGCGCTT CCCCTTCTAC CGCGACAACC CCAAAAAGTG GCAGAACAGC 960
 ATCCGCGACA ACCTTCAACT CAACGACTGC TTCTCTAAGA TCCCGCGCGA GGGCGGCGCG 1020
 CCGGTAAGG GCAACTACTG GCGCTCGAC CCAACCGGG AGGACATGTT CGAGAGCGCG 1080
 AGCTTCTGCG GCGCGCGCAA GCGCTTCAAG CGCTCGGACC TCTCCACTTA CCGCGCTTAC 1140
 20 ATGCAAGAGC CGGCGGCTGC CGCAGCGCGC GCTGCCGCGC CCGCGCGCGC CCTACCGCGG CGCGCTCTAT 1200
 GCGCGCATCT TCCAGGCGCG GGTGCCGCGC GCGCGCGCCC CCTACCGCGG CGCGCTCTAT 1260
 GCAAGGCTAG CGCGCGCGCT GCTGCCGCGC CCGCTCCAG TCTACTACCC CGCGCGCTCG 1320
 CCGGCGCTCT GCGCGCGCTT GCGCTGTGTT CCGTGGCGGC CGCTCAGCCC AGAGCTGGGG 1380
 CCGCACCGCT CGGCGCGCGG GCGCTCTGCG GCGCTTGTCT CCGCGCGCGC CCGCGCTTAC 1440
 25 ACCACCGGCT ACCAGCGCGC AGGCTGCACC GCGCGCGCGC GCGCCAAACC CTCTGCTAT 1500
 GCGGCTGCTT ACAGCGCGCC CGACGCGCGC TACCGCGAGG GCGCGCGCAG TGCGATCTTT 1560
 GCGCGCTGCT GCGCGCTGCG GGGACCGCGT TCGCCCCCAG CCGCGCGCAG CAGTGGCGCG 1620
 GTGGAGACCA CGGTGACTCT CTACGCGCGC ACCTCGCGCG GCGCACTTGG AGCGCTGGGA 1680
 GCGTCTGACA ACCTGCGCGG GCACTCGGA GGGGCCAGTG CAGCGCGCTA CCACTGCTCG 1740
 30 CATGCTGCGC CTATATCCCG TGGGATAGAT CGGTTCGTGT CCGCATGTGT AGCCAGCGTA 1800
 GGGACGAAAA CTCATAGACA CATCGGCTGT TCACACGTTT CCGCGAACCT GAGAACGAA 1860
 AGGAATGAGG AGAGGACTCA ACTGGGACCC ACCTGGGAAA GACCGAGCAG GCCACAGAGG 1920
 CTCGCTCTCC CCGCGCACAG CGTAGGCACC CTGTGTAAT TGTAACCGGG AGGAGGTGGG 1980
 GCGAGGACAG CAGAGCCCTT GGAATGACAG AGGAGCCCTC GATGGAGCGA AGCCCTCAAA 2040
 35 CGGGATGCTT TCTGGCATTC TATCGGGGAG GGTCTTGGC GGTAAACAGA GGGCAGCGTA 2100
 GTGTCAACAC CAGAGACGAG GATCCAAATT GTGGGGAATC AGTTTCAGCC TTCCATGTGC 2160
 TGCGGAACT CCGGCGCTTT TACCGGTTTC GTCTCTAGT GCCTTTAACT GCGTTACTAC 2220
 AATAAAGGCG TGGCGCAGCG CCTTCTTCT TAAAGTGAGG AGGACAAATT TGCAAAAGAA 2280
 ATAGGCTTTT CTCTTTTCTT AAATTTGAGA AATCTCTGCT CTGGTTGACC TGGGCTGGTT 2340
 40 TTCCCTGTCT CTGAGAACTT GAGACCTAGC TCCGAGTTGA ACTGTGCTC AGCACTCCAG 2400
 TCCCATCAC TGAACCTTCA GTCTCCCGCA TCTGTACAC TAGAGGGCTG CAGGACTCTA 2460
 TCCACCGCCC CCGGGTTATC ATTGAGGGCC CCATCATCTT GGATGTCTGC CTGCGTATTT 2520
 GGCAGCAATG GTGGGCCACC CAGGCGCTCT GAGTAGCCAC CCAAGCCCTA GCGCTGTGTT 2580
 TAGGGACCGG AAAAGAGTTC ATGGCCAGGC GTCTAACCTA AAGTCCCGAG ATTGGCTCCA 2640
 45 GGCAGCAATT ATATCATTAAC TTATTGAAT TTTGAGCAGG ACCTGTGCTT AATTTCATGG 2700
 CTGTTACTGC CCACTCATAA ATCTGCTTTT CCATTATAAG GCAGAGAGAA GTACATTCTG 2760
 TCATTTGTCC ACTGTTTCTT GTCATCAGCG AGCCCTGGAC CCAAGGGGTG AACTAAAGTT 2820
 TAAGGAGATG AGAGGATTCA AGGAGCCCGT TGGTGACGCG TTTCAGTAGC TGGGAGGGCG 2880
 TCTTCCATCC CCAAGACCCC CTGCTACACC TCAGCAGCCT CCCCCTATGA AAAAGGAAAG 2940
 50 AGAAAAATTA AGTTAGGSCA GTCAATAAG TGAGCTTTAG AAAGAACTG GAATTTTAA 3000
 TTCAATTTGT ATCTTGCTTA AGTAGCAGGC TCACTAAAAT TAGAGAAAGT CCAATAAATC 3060
 TCCCGCTTTC CCTTGAGAAA TCTTTAAGTT TCGATTCTGG AGCAAAAATC TTCAGCATT 3120
 AATAATTTAG AGGCTCCATT CACAGCTTTC AGATAAACTG GAGTGTTCAG ATGGACTGTT 3180
 TTAATAAAAA TCTTTGAGCA AGTGAGTTAT GGCAAGAGAA ACTCAGCCTC TTTCTGTATA 3240
 55 AACTTAACAG GGAAGGCGTG GGTGTGAAA AAGAAGATTG TATGAAAACC ATTGGTAATT 3300
 TTTATTTTAT ATTTTGGGGA CTGCACTATC CTGTTACAGA AGACATGTGA ACTTGGTTCA 3360
 GTCCAAATGG GGAATTTGAT AAACCACTGC TCTCCATTAG AATATGGTG CAAGCCACAT 3420
 ATGTAATTTT AAATATTCTA GTAGCCACAT TAATAAAGT AAAAGAAACA AAAAAAAAAA 3480
 AA

Seq ID NO: 317 Protein sequence:
Protein Accession #: NP_004464

1 11 21 31 41 51
 65 FKHLTHYROI DTRANSCRIP TIONFACTOR TTFMTAESGP PPPQPEVLAT VKERGETAA 60
 GAGVPGCATG RGAGGRRRKR PLQRGKPPYS YIALIAMATA HAPERRLTGL GIYKFITERP 120
 PFFYRDNPKKW QNSIRHNLTL NDCFLKIPRE AGRPGKGNWY ALDPNAEDMF ESGSFLRRRK 180
 RFRKSDLSY PAYMHDAAA AAAA AAAAIFPGA VPAARPPYFG AVYAGYAPPS 240
 70 LAAPPVVPYV AASPFGPCRVF GLVPERPLSP ELGPAPSGPG GSCAFASAGA PATTTGYQPA 300
 GCTGARPNP SAYAAAYAGP DGAYPQAGS AIFAAAGRLA GPASPPAGGS SGGVETTVDF 360
 YGRTSPGQFG ALGACYNPGG QLGGASAGAY HARHAAAYPG GIDRFVSAM

Seq ID NO: 318 DNA sequence
Nucleic Acid Accession #: NM_005688
Coding sequence: 126..4439

1 11 21 31 41 51
 80 CCGGGCAGGT GGCTCATGCT CGGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCTGGAGC 60
 AGGGGCGCAG GAATTCCTGAT GTGAACTAA CAGTCTGTGA GCCCTGGAAC CTCGCTCAG 120
 AGAAGATGAA GGATATCGAC ATAGGAAAAG AGTATATCAT CCCCAGTCTT GGGTATAGAA 180
 GTGTGAGGGA GAGAACCAGC ACTTCTGGGA CGCACAGAGA CCGTGAAGAT TCCAAGTTCA 240
 85 GGAGAACTCG ACCGTGGGAA TGCCAAGATG CCTTGGAAAC AGCAGCCCGA CCGAGGGGCC 300
 TCTCTCTTGA TGCTCCATG CATTCTCAG TCAGAACTCT GGATGAGGAG CATCCCAAGG 360
 GAAAGTACCA TCATGGCTTG AGTGCTCTGA AGCCCATCG GACTACTTCC AAACACCAGC 420
 ACCCAGTGA CAATGCTGGG CTTTTTCTCT GTATGACTTT TCGTGGCTT TCTTCTCTGG 480

	CCCGTGTGGC	CCACAAGAAG	GGGGAGCTCT	CAATGGAAGA	CGTGTGGTCT	CTGTCCAAGC	540
	ACGAGTCTTC	TGACGTGAAC	TGCAGAAGAC	TAGAGAGACT	GTGGCAAGAA	GAGCTGAATG	600
	AAGTTGGGCC	AGACGCTGCT	TCCTTGGGAA	GGGTTGTGTG	GATCTTCTGC	CGCAACAGGC	660
5	TCATCTGTGC	CATCGTGTGC	CTGATGATCA	CGCAGCTGGC	TGGCTTCAGT	GGACAGGCTT	720
	TCATGGTGAA	ACACCTCTTG	GAGTATACCC	AGGCAACAGA	GTCTAACCTG	CAGTACAGCT	780
	TGTTGTAGT	CTCGGGCTTC	CTCCTGACGG	AAATCGTGCG	GTCTTGGTCG	CTTGCACTGA	840
	CTTGGGCATT	GAATTACCGA	ACCGGTGTCC	GCTTGGCGGG	GGCCATCTTA	ACCATTGGCAT	900
	TTAAGAAGAT	CCTTAAGTTA	AAGAACATTA	AAGAGAAATC	CCTGGGTGAG	CTCATCAACA	960
10	TTTGCTCCAA	CGATGGGCAG	AGAATGTTTG	AGGCAGCAGC	CGTTGGCAGC	CTGTGGGCTG	1020
	GAGGACCCGT	TGTTGCCATC	TTAGGCAATG	TTTATAATGT	AATTATTCTG	GGACCAACAG	1080
	GCTTCTGGG	ATCAGCTGTT	TTTATCCTCT	TTTACCAGC	AATGATGTTT	GCATCAGGCG	1140
	TCACAGCATA	TTTCAGGAGA	AAATGCGTGG	CGGCCACGGA	TGAACGTGTC	CAGAAGATGA	1200
	ATGAAGTCT	TACTTACATT	AAATTTATCA	AAATGTATGC	CTGGGTCAAA	GCATTTCTCT	1260
	AGAGTGTTC	AAAAATCCCG	GAGGAGGAGC	GTGGATATT	GGAAAAAGCC	GGGTACTTCC	1320
15	AGGGTATCAC	TGTGGGTGTG	GCTCCCATTT	TGGTGGTGAT	TGCCAGCGTG	GTGACCTTCT	1380
	CTGTTTATAT	GACCTGGGGC	TTCCATCTGA	CAGCAGCACA	GGCTTTTACA	GTGGTGACAG	1440
	TCCTTCAATT	CATGACTTTT	GCTTTGAAG	TAACACCGTT	TTCACTAAAG	TCCCTCTCAG	1500
	AAGCTCAGT	GCTGTGTGAC	AGATTTAAGA	GTTTGTTCCT	AATGGAAGAG	GTTTCACTGA	1560
20	TAAAGAACAA	ACCAGCCAGT	CCTCACATCA	AGATAGAGAT	GAATAATGCC	ACCTTGGCAT	1620
	GGGACTCCTC	CCATCTCCAGT	ATCCAGAACT	CGCCCAAGCT	GACCCCAAAA	ATGAAAAAAG	1680
	ACAAGAGAGG	ATCCAGGGGC	AAGAAAGAGA	AGGTGAGGCA	GCTGCAGCGC	ACTGAGCATC	1740
	AGGCGGTGCT	GGCAGAGCAG	AAAGGCCACC	TCCTCCTGGA	CAGTGACGAG	CGGCCAGCTC	1800
	CGGAAGAGGA	AGAAGGCCAG	CACATCCACC	TGGGCCACCT	CGCTTACAG	AGGACACTGC	1860
25	ACAGCATCGA	TCTGGAGATC	CAAGAGGGTA	AATCGTGGG	AATCTGCGGC	AGTGTGGGAA	1920
	GTGGAAAAAC	CTCTCTCATT	TCAGCCATTT	TAGGCCAGAT	GACGCTTCTA	GAGGGCAGCA	1980
	TTGCAATCAG	TGCAACCTTC	GCTTATGTGG	CCCAGCAGGC	CTGGATCCTC	AATGCTACTC	2040
	TGAGAGACAA	CATCCTGTTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
	ACAGCTGCTG	CTGAGGCCCT	GACCTGGCCA	TTCTTCCAG	CAGCGACCTG	ACGGAGATTT	2160
30	GAGAGCGAGG	AGCCAACTTG	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCTT	GGCCGGGCTT	2220
	TGTATAGTGA	CAGGAGCATC	TACATCCTGG	ACGACCCCTT	CAGTGCCTTA	GATGCCCATG	2280
	TGGGCAACCA	CATCTTCAAT	AGTGTATATC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
	TTGTTACCCA	CCAGTTACAG	TACCTGGTTG	ACTGTGATGA	AGTGTCTTTC	ATGAAAAGAG	2400
	GCTGTATTAC	GGAAAGAGGC	ACCCATGAGG	AACGTGATGA	TTTAAATGGT	GACTATGCTA	2460
35	CCATTTTAA	TAACTGTGTT	CTGGGAGAGA	CACCGCCAGT	TGAGATCAAT	TCAAAAAGAG	2520
	AAACCACTGG	TTACAGAAAG	AAGTCACAAG	ACAAGGGTCC	TAAACAGGGA	TCAGTAAAGA	2580
	AGGAAAAAGC	AGTAAAGCCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
	GTTCACTGCC	CTGGTCAGTA	TATGGTGTCT	ACATCCAGGC	TGCTGGGGGC	CCCTTGGCAT	2700
	TCCTGGTATT	TATGGCCCTT	TTCTATGCTG	ATGTAGGCAG	CACCGCCTTC	AGCACTGGTT	2760
40	GGTTGAGTTA	CTGGATCAAG	CAAGGAAGCG	GGAACACCA	TGTGACTCGA	GGGAACGAGA	2820
	CCTCGGTGAG	TGACAGCATG	AAGGACAATC	CTCATATGCA	GTACTATGCC	AGCATCTACG	2880
	CCCTCTCCAT	GGCAGTCATG	CTGATCCTGA	AAGCCATTGG	AGGAGTTGTC	TTTGTCAAGG	2940
	GCACGCTGCG	AGCTTCTCCT	CGGCTGCATG	ACGAGCTTTT	CCGAAGGATC	CTTCCGAGCC	3000
	CTATGAAGTT	TTTGAACAGC	ACCCCAACAG	GGAGGATTCT	CAACAGGTTT	TCCAAAGACA	3060
45	TGGATGAAGT	TGACGTGCGG	CTGCCGTTC	AGGCCGAGAT	GTTTATCCAG	AACGTTATCC	3120
	TGGTGTCTCT	CTGTGTGGGA	ATGATCGCAG	GAGTCTTCCC	GTGGTTCCTT	GTGGCAGTGG	3180
	GGCCCTCTGT	CATCCTCTTT	TCAGTCTGTC	ACATTGTCTC	CAGGTCCTGT	ATTCCGGAGC	3240
	TGAAGCGTCT	GGACAATATC	ACGCACTCAC	CTTCTCTCTC	CCACATCAGC	TCCAGCATAC	3300
	AGGGCCTTGC	CACCATCCAC	GCCTACAATA	AAGGGCAGGA	GTTTCTGCAC	AGATACCAAG	3360
50	AGCTGCTGGA	TGACAACCAA	GCTCCTTTT	TTTGTGTTAC	GTGTGCGATG	CGGTGGCTGG	3420
	CTGTGCGGCT	GGACCTCATC	AGCATCGCCC	TCATCACCAC	CACGGGCTGT	ATGATCGTTC	3480
	TTATGCAACG	GCAGATTCCC	CCAGCCTATG	CGGGTCTGCG	CATCTCTTAT	GCTGTCCAGT	3540
	TAAAGGGGCT	GTCCAGGTTT	ACGGTCAGAC	TGGCATCTGA	GACAGAAGCT	CGATTCACTT	3600
	CGGTGAGAG	GATCAATCAC	TACATTAAGA	CTCTGTCTCT	GGAAAGCAGC	GCCAGAAATTA	3660
55	AGAACAAGGC	TCCCTCCCTT	GACTGGCCCC	AGGAGGGAGA	GGTGACCTTT	GAGAACGCGAG	3720
	AGATGAGGTA	CCGAGAAAAA	CTCCCTCTTG	TCCTAAAGAA	AGTATCCTTC	ACGATCAAAC	3780
	CTAAAGAGAA	GATTGGCAAT	GTGGGGCGGA	CAGGATCAGG	GAAGTCTCTG	CTGGGGATGG	3840
	CCCTCTTCCG	TCTGGTGGAG	TTATCTGGAG	GCTGCATCAA	GATTGATGGA	GTGAGATACA	3900
	GTGATATTGG	CTTGGCGGAC	CTCCGAAGCA	AATCTCTAT	CATTCTCTCA	GAGCCGGTGC	3960
60	TGTTCACTGG	CAGTGTGAGA	TCAAATTGAG	ACCCCTTCAA	CCAGTACACT	GAAGACACGA	4020
	TTTGGGATGC	CCTGGAGAGG	ACACACATGA	AAGAATGTAT	TGCTCAGCTA	CCTCTGAAGC	4080
	TTGAATCTGA	AGTGTGAGAG	AATGGGGATA	ACTTCTCAGT	GGGGGAACGG	CAGCTCTTGT	4140
	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTAGATGAA	GCCACAGCTG	4200
	CCATGGACAC	AGAGACAGAC	TTATTGATTC	AAGAGACCAT	CCGAGAAGCA	TTTGACAGCT	4260
65	GTACCATGCT	GACCATGTCC	CATCGCCTGC	ACACGGTTCT	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGCCCA	GGGACAGGTG	GTGGAGTTTG	ACACCCCATC	GGTCTTCTGT	TCCAACGACA	4380
	GTTCCCGATT	CTATGCCATG	TTTGTGCTG	CAGAGAACAA	GGTGCCTGTC	AAGGGCTGAC	4440
	TCCTCCCTGT	TGACGAAGTC	TCTTTTCTTT	AGAGCATTGC	CATTCCCTGC	CTGGGGGGGG	4500
	CCCTCATGCT	GCTCCTCTTA	CCGAAACCTT	GCCTTCTCTG	ATTTTATCTT	TCGCACAGCA	4560
70	GTTCCGGATT	GGCTTGTGTG	TTTCACTTTT	AGGGAGAGTC	ATATTTTGAT	TATTGTATTT	4620
	ATTCATATAT	CATGTAAACA	AAATTTAGTT	TTTGTCTCTA	ATTGCACTCT	AAAAGGTTCA	4680
	GGGAACCGTT	ATTATAATTG	TATCAGAGGC	CTATAATGAA	GCTTTATACG	GTAGCTATA	4740
	TCTATATATA	ATTCTGTACA	TAGCCTATAT	TTACAGTGAA	AATGTAAAGT	GTTTATTTTA	4800
	TATTAAATA	AGCACTGTGC	TAATAACAGT	GCATAATCCT	TTCTATCAIT	TTTGACAGT	4860
75	TTGCTGTACT	AGAGATCTGG	TTTGTCTATT	AGACTGTAGG	AAGAGTAGCA	TTTCACTCTT	4920
	CTCTAGCTGG	TGGTTTCACG	GTGCCAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTGTGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCTCTGCCC	GCCTCCCCAC	AGCCGCTCCA	GGGGTGGCTG	5040
	GAGACGGGTG	GGCGGCTGGA	GACCATGCAG	AGCCCGGTGA	GTTCTCAGGG	CTCCTGCTCT	5100
	CTGTCTCTGT	GTCATCTTAT	GTTTCTGTCA	GGGAGAGCAG	GGGGGGAAGC	CCAGGCCCTT	5160
80	TTTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTGCTCT	TCTTCTTTTT	GCTGTGTGTT	CTAAACAGAA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCACTGCC	TCAGGTTTCT	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACCT	5340
	GTTGGTTCCA	AGCCCTGGAG	CCAATGCTGT	CTTTTGTAGG	TGGCACTTTT	TCAATTCCTT	5400
	ATTCACACAC	CTCCACAGTT	CAGTGGCAGG	GCTCAGGATT	TGGTGGGTCT	GTTTCTCTTT	5460
85	CTCACCGCAG	TGCTGCGACA	GTCTCTCTCT	CTCTCTCCCC	TCAAAGTCTG	CAACTTTAAG	5520
	CAGCTCTTGC	TAACTAGTGT	CTCACACTGG	CGTAGAAGTT	TTTGTACTGT	AAAGAGACCT	5580
	ACCTCAGGTT	GCTGGTGTCT	GTGTGTTTCC	GCAAAACCCC	TTTGTGCTGT	TTTGTGCTGT	5640
	GGGGCTGGTA	GCTCAGGTGG	GCGTGGTCAC	TGCTGTCTAT	AGTTGAATGG	TCAGCTGTGC	5700

ATGTCGTGAC CAACTAGACA TTCTGTCGCC TTAGCATGTT TGCTGAACAC CTTGTGGAAG 5760
 CAAAAATCTG AAAATGTGAA TAAATATT TTGGATTITG TAAAAAATAA AAAAAAATAA 5820
 AAAAAAATAA AAAAAAATAA

Seq ID NO: 319 Protein sequence:
 Protein Accession #: NP_005679

1 11 21 31 41 51
 MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRF TRPLECQDAL ETAARAEGLS 60
 LDASMSQLR ILDEEHPKKG YHGLSALKP IRTTSKHQHP VDNAGLFSCM TFSWLSLAR 120
 VAHKGELSM EDVWSLSKHE SSDVNCRRLE RLWQEELNEV GPDAASLRRV VWIFCRTRI 180
 LSVICLMTIQ LAGFSGPAFM VKHLLLEYQA TESNLQYSL LVLGLLLEI VRSWSLALTW 240
 ALNYRTGVR LRGAILTMAPK KILKLNKKE KSLGELINIC SMDGORMFEA AAVGSLLAGG 300
 PVVAILGMY NVIILGPTGF LGSVAFILFY PAMMFASRLT AYFRKCVAA TDERVQKME 360
 VLYIKFIKM YAWVKAFSQS VQKIREEERR ILEKAGYFQG ITVGVAPIVV VIASVVTFSV 420
 HMTLGDFTA QAFTVTTFV NSMTFALKVT PFSVKSLSEA SVAVDRFKSL FMEEVHMIL 480
 NKPASPHIKI EMKNATLAWD SSHSSIQNSP KLTPKMKDKK RASRGKKEKV RQLQTEHQA 540
 VLAEQKGLL LDSDERPSP EEEGKHILG HLRLQRTLHS IDLEIQEGKL VGIQSVGSG 600
 KTSLSIALIG QMILLEGSLA ISGTFAVVAQ QAWILNATLE DNILFGKEYD EERYNSVLNS 660
 CCLRPDLALF PSSDLTEIGE RGNALSGGQR QRISLARALY SDRSIYILDD PLSALDAHV 720
 NHIFNSALIK HLKSKTVLVF THQLQYLVDG DEVIFMKEGC ITERGTHEEL MNLGNDYATI 780
 FNNLLGETP PVEINSKKEP GSGQKKSQDK GPKTGSVKKE KAVKPEEGQL VQLEEKQGS 840
 VPWSVYGVYI QAAGGPAFL VIMALFMINV GSTAFSTWML SYWIKQSGSN TTVTRGNETS 900
 VSDSMKDNPH MQYASIIYAL SMAMVLLILKA IRGVVFKGT LRASSRLHDE LFRILRSPM 960
 KPFDTTPTGR ILNRFSKMD EVDVRLPFQA EMFIQNVILV PFCVGMIAGV FFWFLVAVGP 1020
 LVILFSLVLI VSRVLIRELK RLDNITQSPF LSHITSSIQG LATIHAYNKG QFLHRYQEL 1080
 LDDNQAPPFL PTCAMRWLAV RLDLISIALI TTTGLMIVLM HQQIPPAYAG LAISYAVQLT 1140
 GLFQFTVRLA SETEARTTSV ERINHYIKL SLEAPARIKN KAPSPDWQE GEVTFENAE 1200
 RYRENLPVLV KKVSPFTKPK EKIGIVGRGT SGKSSLGML FRLVELSGGC IKIDGVRIS 1260
 IGLADLRSLI SIIPQEPVLF SGTVRSNLDP FNQYTEDQIW DALERTHME CIAQLPLKLE 1320
 SEVMENGDNF SVGERQLLCC ARALLRHCKI LILDEATAAM DTETDLIQL TIREAFADCT 1380
 MLTIAHRLHT VLGSDRIMVL AQGQVVEFDT PSVLLSNDSS RPYAMFAAAB NKVAVKG

Seq ID NO: 320 DNA sequence
 Nucleic Acid Accession #: AK022089.1
 Coding sequence: 181-1488

1 11 21 31 41 51
 AGCAGTGTCA CAACTTCCAG CAACTTCTC AGCCGGCTAC TAATGAGCTG AAAGCCAGGA 60
 ACATCCGAGG AGAAGAGAAA GCTTCCAGCC CTCTCCCTT CACCTGGAA ATCCAGACAC 120
 CCCACCCCCC ACCCTCAGAT CACTTTAAGA TAATTTCTTT ATTGTTTGC CGACAGACCC 180
 ATGGCTCCCT TTGAAGAGAA CTGCTAAAG ACTCGGCATA AAAACAGATC TCCAACTAAA 240
 GACATGGATT CAGAAGAGAA GGAAATTGTG GTTTGGGTTT GCCAAGAAGA GAAGCTTGTC 300
 TGTGGGCTGA CTAAACGCAC CACCTCTGCT GATGTCATCC AGGCTTTGCT TGAGGAACAT 360
 GAGGCTACGT TTGAGAGAAA ACGATTCTT CTGGGGAAGC CAGTGATTA CTGCATCATA 420
 GAGAAGTGA GAGGCTCCGA AAGGGTCTT CTCCACTAA CTAGAATCCT GAAGCTTTGG 480
 AAAGCGTGGG GAGATGAGCA GCCCAATATG CAATTTGTTT TGGTTAAAGC AGATGCTTTT 540
 CTCCAGTTC CTTTGTGGG GACAGCTGAA GCCAAATTAG TGCAAAACAC AGAAAAATTG 600
 TGGGAGCTCA GCCAGCAAAA CTACATGAAG ACTTTACCAC CAGATAAACA AAAAAGAATA 660
 GTCAGGAAAA CTTTCCGAAA ACTGCTAAA ATTAAGCAGG ACACAGTTTC TCATGATCGA 720
 GATAATATGG AGACATTAGT TCATCTGATC ATTTCCAGG ACCATACTAT TCATCAGCAA 780
 GTCAGAGAAA TGAAAGAGCT GGATCTGGAA ATTGAAAAGT GTGAAGCTAA GTTCCATCTT 840
 GATCGAGTAG AAAATGATGG AGAAAACTAT GTTCAGGATG CATATTTAAT GCCCAGTTTC 900
 AGTGAAGTGG AGCAAAATCT AGACTTGCAG TATGAGGAAA ACCAGACTCT GGAGGACCTG 960
 AGCGAAAGTG ATGGAATTGA ACAGCTGGAA GAACGACTGA AATATTACCG AATACTCATT 1020
 GATAAGCTCT CTGCTGAAAT AGAAAAAGAG GTAAAAAGTG TTTGCATTGA TATAAATGAA 1080
 GATGCGGAAG GGGAGAGCTG AAGTGAAGCT GAAAGCTCTA ATTTAGAGAG TGTTAAGTGT 1140
 GATTTTGAGA AAAGCATGAA AGCTGGTTTG AAAATTCACT CTCATTTGAG TGGCATCCAG 1200
 AAAGAGTTA AATACAGTGA CTCATTGCTT CAGATGAAAG CAAAAGAATA TGAATCTCTG 1260
 GCCAAGGAAT TCAATCACT TCACATTAGC AACAAGATG GGTGCCAGTT AAAGGAAAC 1320
 AGAGCGAAGG AATCTGAGGT TOCCAGTAGC AATGGGGAGA TTCTCCCTT TACTCAAGA 1380
 GTATTTAGCA ATTACACAAA TGACACAGAC TCGGACACTG GTATCAGTTC TAACCAAGT 1440
 CAGGACTCCG AAACAACAGT AGGAGATGTG GTGCTGTTGT CAACATAGTT CCAATGGCTC 1500
 CTTTCTGACC TGCTTTCATG TTTTAATGTT TGTTTAATTT AATAGGAAAC CTCATTTTAA 1560
 ATATAACACT CAAAAAATG TAAATCATAT TGTAGTATTC AATAGTTAAT AAAAAGCTCGA 1620
 GAAATGTGTT GTTTCTG

Seq ID NO: 321 Protein sequence:
 Protein Accession #: NP_005438.1

1 11 21 31 41 51
 MAPFGRNLLK TRHKNRSPK DMDSEEKIV VWVCQBEKLV CGLTKRTTSA DVIQALLEEH 60
 EATFGEKRF LKGPSDYCI EKWRGSEVL PPLTRILKLW KAWGDEQPNM QFVLVKADAF 120
 LPVPLWRTAE AKLVQNTKEL WELSPANVMK TLPDPKQKRI VRKTFKRLAK IKQDVTSHDR 180
 DNMTLVHLI ISQDHTIHQQ VKRMKELDL IEKCEAKPHL DRVENDGENY VQDAYLMPSF 240
 SEVEQNLDLQ YEENQTLDDL SESDGIQLE ERLKYRILI DKLSAEIEKE VKSVCIDINE 300
 DAGGAASEL ESSNLESVKC DLEKSMKAGL KIHSHLSGIQ KEIKYSDSL QMKAKEYELL 360
 AKEPNLSLHIS NKDGQCKEN RAKESEVPSS NGEIPPTQR VFSNYTNDTD SDTGISSNHS 420
 QDSEITVGDV VLLST

Seq ID NO: 322 DNA sequence
 Nucleic Acid Accession #: NM_030920.1

Coding sequence: 317-1123

	1	11	21	31	41	51	
5	AGCATTGAAG	GGGAAGGAAC	TGCGGGTGTG	GTGTTGTGTAT	GTGTTGTGTGT	ATGTTGTGTGC	60
	GGCGCGTGGC	TGCGTGTGTG	TGCGCGCGCT	AGTGTGTGGA	CAAGGAGGTG	GGGGCAGCTG	120
	AGTTAGAGTC	CCAACCTCTG	GACTCCATTT	GCTATTCTCT	TCTTTCTCCC	CCACACCTAT	180
	CTGGTGGTGG	TAGTGGGCGT	TTATATTTGC	GTTCCTTTTC	ATTCTATTCT	AAATCTCTTA	240
	AAAATTTTGG	GTGTTGGGTA	TTGGGGAAGG	CAGGAAAGGG	AAAAGGAGAG	TAGTAGCTGA	300
10	AGAGCAAGAG	GAGGACATGG	AGATGAAGAA	GAAGATTAAC	CTGGAGTTAA	GGAAACAGATC	360
	CCCGGAGGAG	GTGACAGAGT	TAGTCTCTGA	TAATTCGCTG	TGTGTCAATG	GGGAAATTGA	420
	AGGCGTGAAT	GATACCTTCA	AAGAACTAGA	ATTTCTGAGT	ATGGCTAATG	TGGAACTAAG	480
	TTGCGTGGCC	CGGCTTCCCA	GCTTAAATAA	ACTTCGAAAA	TTGGAGCTTA	GTGATAATAT	540
	AATTTCTGGA	GGCTTGGGAG	TCCTGGCAGA	GAATGTCCCA	AATCTTACCT	ACCTCAATCT	600
15	GAGTGGAAAC	AAAATAAAG	ATCTCAGTAC	AGTAGAAGCT	CTGCAAAATC	TTAAAAATTT	660
	GAAAAAGTCT	GACCTGTTTA	ACTGTGAGAT	CACAAACCTG	GAAGATTATA	GAGAAAGTAT	720
	TTTTGAAGTA	CTGCAGCAAA	TCACATACCT	AGATGGATTT	GATCAGGAGG	ATAATGAAGC	780
	GCGGAGCTCT	GAGAGGAGG	ATGATGAGGA	TGGAGATGAA	GATGATGAAG	AGGAAGAGGA	840
	AAATGAAGCT	GGTCCACCGG	AAGGATATGA	GGAAGAGGAG	GAGGAAGAGG	AAGAGGAGGA	900
20	TGAGGATGAG	CATGACATGG	AAGATGAAGC	AGGTTGAGAG	TTGGGAGAGG	GAGAGAGGGA	960
	AGTGGGCTCT	TCATACCTAA	TGAAAGAAGA	AATTCAGGAT	GAAGAAGATG	ATGATGACTA	1020
	TGTTGAAGAA	GGGGAAGAAG	AGGAAGAAGA	GGAAGAAGGA	GGTCTTCGAG	GGGAGAAGAG	1080
	GAAACGAGAT	GCTGAAGAG	ATGGAGAGGA	AGAAGATGAC	TAGATCATTC	TAAGACCAGA	1140
	TTCTCTAATG	TTTCTGGGTG	TGCAATAGAG	TGATCACAAT	TTTGTTCCTT	CATGTACGAT	1200
25	AGCTATCCCT	ACAGAAAGAT	ATGTGTAAC	TTTATAGGA	AAAGTGTGGT	TTTACTATTT	1260
	TTGCCCTATC	ATTCCTAATA	AGAACTAGTC	TGTTAATGAT	CATATTGTAT	GTAGAGAAAA	1320
	ATTTTCATTT	ACTCCCATTT	TGGAATTTCC	TAGCAATTTA	TTTACACTTA	ATTTTTTAAA	1380
	TTCAAGCTTA	CTGATGATGG	CATTTTATAG	CCATAATTTA	AACATGATCA	CTTTTAAACA	1440
	GGTGTAGTAT	GGTGCATTTT	ATTCCTTATT	TATAGATTAA	CTGAAATTTA	AGTTTGTCTAT	1500
30	AATATAAAAT	GACATATAGT	TCTTGAGTGG	TAAGTTGGTT	ATTTTITTAG	AGGTGATCCA	1560
	GGAAATCTTT	GTGTTGAAGG	AGTTACCTTT	TTTTTTTTTT	TTTTTTTTTG	ACTAAGAGTG	1620
	TTTGGTGTCT	TTTTTGTGAC	AAGTAACCTG	GAAATAGAAA	GCAGAATAGT	AAAGGTTCTA	1680
	TTGAGCAACA	TAGTTCATGG	ATTTGTGGA	GGTCTATTC	AGTAATATGG	TTGATGGATT	1740
	TAGTGTGAC	TGATAGAGAT	TTATTTTGA	AGGAAAAATT	GCTTATACTA	AGTCCAGAGA	1800
35	CATGCAGGTG	AGCCCTTTTG	TCAGGCTGCA	AATCATGACA	TGCCGATGGT	TGTTTATTTT	1860
	GTGTTTAGGT	GTGCATTTCT	TTTCTTCTTA	GCAATTCCTT	TATGATCACC	TTCCCTTCTT	1920
	GTTCCTCTCC	CTCCCGCTCT	CTCAAAAGGA	ACTTGGGAAA	CTGTGAAAC	CCAGAAAAAC	1980
	CTTTAGTCTT	ATACCTCAAC	TACGTTTCAG	TCTCTCTGG	GTGTTTAAATA	AGTGAAGTAG	2040
	AAGAAATTGA	GTATTTTCTG	ACATAAGAAT	ATATATCAAA	TACAGTTTAA	TGCAATGAAG	2100
40	TCTCCTTACC	ATAAATGTTT	CTTGGTTGAC	AACATCTAAG	ACAATATTAG	TGGGATGAAG	2160
	AAAGAAAAGC	AGGGGTGCTT	TTGGAAGCAG	TGTTAGTGTT	CCTCAAAAGT	CGGAACAATT	2220
	GCCTGTGAT	ATATTAATAA	GACATTAAG	TCAAAATTTA	ATGTTGGCCT	CTCAATGAT	2280
	TTGGATACCA	CTCTGCAAG	TATTTCTAAC	CTTTAATTTT	CAGTTTAAAT	ACAGATATAA	2340
	TAATAGCATT	TAATTTGAAT	ATACTAGGCA	GCTGGAAGAG	TATTTGAAAC	TAAATTGACA	2400
45	TTAAATTAAT	GATTTGTTTT	CAAGTGGATG	TCCATTAAAA	GTAGAAAAAT	ATTTGGGATA	2460
	AGTGAGTGTG	TGTTTCTCTA	CATGGCTACT	AAATAAAATA	TAATGAGTAT	ACAAGTATAT	2520
	CTCCTCTTTT	GCTATGAGG	CTCCATGTTT	AAGGCAATGG	CTTTTAAAT	CTTGGCTATC	2580
	TAAATTTTGT	TCCCTTTGTT	TTGAATATTT	GTAAGTTTIT	AAGAAGTTAG	TGTCAGCAAA	2640
	TTAATTTGAAG	TTATGCTTCT	ATACTGGGAC	ATATTTAAAT	ACTGAGTATA	GTACTGTCTG	2700
50	TACTGCTTCT	ACAATGTAAA	ATGTATGACT	TGGTGTITTA	AAGTAAAAAT	TATGATGTTA	2760
	CTGTGTGAGA	AACTAAAAAT	GTTGTACAAC	TGACCGAAG	AAAACCCCTG	GGGATAAGTT	2820
	TAGTGAGGGG	ATTGGAATCC	CCAAAAAGAT	AACATTTTTC	TTCTGCTTTT	AAAAACTGAA	2880
	ATTCCCTGTT	CTAGTTCTTA	ACAATTTCTA	TTACATACTA	TGCCAGATTA	CAAAATACCT	2940
	ATTTTAAAAA	TGAATCTAT	ATATTGACTT	TCTTATCAAT	CATCTTACTG	TGCAATCAAA	3000
55	ATTAGAGTAC	TTTGGTTTGA	AAACAACACT	TAGAGCCTCC	AGATAACTTT	TAAGACTTAT	3060
	TTAGCTTTGT	GGGTGGTATT	TTTATGCAAA	TAAAGTAAGG	TGGGTTTAT	ATTTGTAGTA	3120
	AGTTTTCGGT	CCTATTTTAA	TGCTCTTTGT	ATGGCAGTAT	GTATATATTG	TGTTAAGTTC	3180
	CTCAAGAAAT	TCCTTAAAAA	CTTTGAAGTT	AATACTTTTG	TGCAACTGTG	TTTTGAATAA	3240
60	AGCCATGACA	GTGTTAAAAA	CAAAAC				

Seq ID NO: 323 Protein sequence:
Protein Accession #: NP_112182.1

	1	11	21	31	41	51	
65	MEMKKKINLE	LRNRSPEEVT	ELVLDNCLCV	NGEIEGLNDT	FKELEFLSMA	NVELSSLARL	60
	PSLNKLRLLE	LSDNIISSGL	EVLAERKCPNL	TYLNLSGNKI	KDLSTVEALQ	NLKNLKSLLD	120
	FNCEITNLED	YRESIFELLQ	QITVLDGFDQ	EDNEAPDSEB	EDDEDGDEDD	EEEREENAGP	180
70	PEGYEEEEEE	EEDEDEDEDE	DEDEAGSELG	EGEEEVGLSY	LMKEEIQDEE	DDDDYVEEGE	240
	EEEEEEEGGL	RGKKRRDAE	DDGEEEDD				

Seq ID NO: 324 DNA sequence
Nucleic Acid Accession #: NM_003812
Coding sequence: 224..2722

	1	11	21	31	41	51	
75	TCCTCTCGGT	CCCCCCCCGG	GAGTGGCTGC	GAGGCTAGGC	GAGCCGGGAA	AGGGGGCGCC	60
	GCCAGGCCCC	GAGCCCCGCG	CCCGTGCCTC	CGAGCCCGGA	GCCCCCTGCC	CGCGCGCGCA	120
80	CCATGCGCGC	CGAGCCGCGC	TGACCGGCTC	CGCCCGCGCG	CGCCCGCGAG	CTAGCCCGGC	180
	GCTCTCGCGC	GCCACAGGGA	GCGGCGCGCG	GGAGCTATGA	GCCATGAAGC	CGCCCGCGAG	240
	CAGCTCGCGC	CAGCCGCCCC	TGGCGGGCTG	CAGCCTTGCC	GGCGCTTCTT	CGGGCCCCCA	300
	ACGCGGCCCC	GCGGCTCGGG	TGCTTGCCAG	CGCCCGCGCC	CGCAAGCGCG	CCTGCGCGCT	360
	GCTTCTCGTC	CTTCTCTGCG	TGCTTCCGCT	CGCCCGCTCG	TCCCGGCCCC	GCGCTCGGGG	420
85	GGCTGCTGCG	CCCAGCGCTC	CGCATTTGAA	TGAAACTGCA	GAAAAAATTT	TGGGAGTCTT	480
	GGCAGATGAA	GACAAATAC	TGCAACAGAA	TAGCAGCAGT	AATATCAGTT	ACAGCAATGC	540
	AATGCAGAAA	GAAATCACAC	TGCCTTCAAG	ACTCATATAT	TACATCAACC	AAGACTCGGA	600

WO 02/086443

PCT/US02/12476

	AAGCCCTTAT	CACGTTCTTG	ACACAAAGGC	AAGACACCAG	CAAAAACATA	ATAAGGCTGT	660
	CCATCTGGCC	CAGGCAAGCT	TCCAGATTGA	AGCCTTCGGC	TCCAAATTCA	TTCTTGACCT	720
	CATACCTGAAC	AATGGTTTGT	TGTCTTCTGA	TTATGTGGAG	ATTCACTACG	AAAAATGGGA	780
5	ACCACAGTAC	TCTAAGGGTG	GAGAGCAGTG	TTACTACCAT	GGAAGCATCA	GAGGCGTCAA	840
	AGACTCCAAG	GTGGCTCTGT	CAACCTGCAA	TGGACTTCAT	GGCATGTTTG	AAGATGATAC	900
	CTTCGTGTAT	ATGATAGAGC	CACCTAGAGCT	GGTTCATGAT	GAGAAAAGCA	CAGGTCGACC	960
	ACATATAATC	CAGAAAACTT	TGGCAGGACA	GTATTCTAAG	CAAAATGAAGA	ATCTCACTAT	1020
	GGAAAGAGGT	GACCAAGTGG	CCTTTCTCTC	TGAATTACAG	TGTTTGAATA	GAAGGAAGAG	1080
	AGCAATGAAT	CCATCAGGTG	GTATATTGTA	AGAAATGAAA	TATTTGGAAC	TTATGATTGT	1140
10	TAATGATCAC	AAAAAGTATA	AGAAGCATCG	CTCTTCTCAT	GCACATACCA	ACAACTTTGC	1200
	AAAGTCCGTG	GTCAACCTTG	TGGATTCTAT	TTACAAGGAG	CAGCTCAACA	CCAGGGTTGT	1260
	CCTGGTGGCT	GTAGAGACCT	GGACTGAGAA	GGATCAGATT	GACATCAGCA	CCAAACCTGT	1320
	GCAGATGCTC	CATGAGTTCT	CAAAATACCG	GCAGCGCATT	AAGCAGCATG	CTGATGCTGT	1380
	GCACCTCATC	TGCGCGGTGA	CATTTCACTA	TAAGAGAAGC	AGTCTGAGTT	ACTTTGGAGG	1440
15	TGCTGTGTTT	CGCACAAGAG	GAGTTGGTGT	GAATGAGTAT	GGTCTTCCAA	TGGCAGTGGC	1500
	ACAAGTATTA	TGCGAGAGCC	TGGCTCAAAA	CCTTGGAAATC	CAATGGGAAC	CTTCTAGCAG	1560
	AAAGGCAAAA	TGTGACTGCA	CAGAACTCTG	GGTGGCTGCG	ATCATGGAGG	AAACAGGGGT	1620
	GTCCCATTTCT	CGAAAAATTTT	CAAAAGTGAG	CATTTGGAGG	TATAGAGACT	TTTTACAGAG	1680
	AGGAGGTGGA	GCCTGCCTTT	TCAACAGGCC	AAACAAAGCTA	TTTGAGCCCA	CGGAATGTGG	1740
20	AAATGGATAC	GTGGGAAGCTG	GGGAGGAGTG	TGATTGTGGT	TTTCATGTGG	AATGCTATGG	1800
	ATTATGCTGT	AAGAAATGTT	CCCTCTCCAA	CGGGGCTCAC	TGCAGCGAGC	GGCCCTGCTG	1860
	TAACAATACC	TGATGTCTTT	TTGAGCCACG	AGGATATGAA	TGCCGGGATG	CTGTGAACGA	1920
	TGCTGATATT	ACTGAATTTT	GTACTGGAGA	CTCTGGTCAG	TGCCCAACCA	ATCTTCATAA	1980
	GCAAGACGGA	TATGCATGCA	ATCAAAATCA	GGGCGCTGCG	TACATGGCGG	AGTGAAGAC	2040
25	CAGAGACAAC	CAGTGTGAGT	ACATCTGGGG	AAACAAAGGCT	GCAGGGTCTG	ACAAGTCTGT	2100
	CTATGAAAAG	CTGAATACAG	AAGGCACTGA	GAAGGGAAGC	TGCGGGAAGG	ATGGAGACCG	2160
	GTGGATTGAG	TGCAGCAAA	ATGATGTGTT	CTGTGGATTC	TTACTCTGTA	CCAATCTTAC	2220
	TGAGCTCCA	CGTATTGCTC	AACCTCAGGG	TGAGATCATT	CCAACCTCTT	TCTACCATCA	2280
	AGGCGGGTG	ATTGACTGCA	GTGGTGCCCA	TGTAGTTTAA	GATGATGATA	CGGATGTGGG	2340
30	CTATGTAGAA	GATGGAAGCG	CATGTGGCCC	GTCTATGATG	TGTTTAGATC	GGAAGTGCCT	2400
	ACAATTCAC	GGCCCTAAAT	TGAGCAGCTG	TCCACTCGAT	TCCAAGGGA	AAGTCTGTTC	2460
	GGGCCATGGG	GTGTGTAGTA	ATGAAGCCAC	CTGCATTTGT	GATTTCACTT	GGGCAGGGAG	2520
	AGATTGCAAT	ATCCGGGATC	CAGTTAGGAA	CCTTCAACCC	CCCAAGGATG	AAGGACCCAA	2580
	GGGTCTTAGT	GGCCCAATC	TCATAATAGG	CTCCATCGCT	GGTGCCATCC	TGGTAGCAGC	2640
35	TATGTCTCTT	GGGGCCACAG	GCTGGGGATT	TAAAAATGTC	AAGAAGAGAA	GGTTCGATCC	2700
	TACTCAGCAA	GGCCCATCTT	GAATCAGCTG	CGCTGGATGG	ACACCGCTCT	GCACGTGTTG	2760
	ATTCTGGGTA	TGACATACCT	GCAGCAGTGT	TACTGGAATC	ATTAAGTTTG	TAAACAAAA	2820
	CTTTGGGTGG	TAATGACTAC	GGAGCTAAAG	TTGGGGTGAC	AAGGATGGGG	TAAAGAAAA	2880
40	CTGTCTCTTT	TGGAATAAT	GTCAAAGAAC	ACCTTTTACC	ACCTGTGAGT	AAACGGGGGA	2940
	GGGGGCAAAA	GACCATGCTA	TAAAAAGAAC	TGTTCCAGAA	TCTTTTTTTT	TCCCTAATGG	3000
	ACGAAGGAAC	AACACACACA	CAAAAATTAA	ATGCAATAAA	GGAATCATTAA	AAAA	

Seq ID NO: 325 Protein sequence:
Protein Accession #: NP_003803

45	1	11	21	31	41	51	
	MKPPSSSRQ	PPLAGCSLAG	ASCGPQRGPA	GSVPASAPAR	TPPCRLLLV	LLLPLLAASS	60
50	RPRANGAAR	SAPHWNETA	KNLGVLADED	NTLQNNSSN	ISYSNAMQKE	ITLPSRLIYY	120
	INQDSSEPYH	VLDTKARHQ	KHNKAVHLAQ	ASFQIEAFPS	KFILDILINN	GLLSSDYVEI	180
	HYENGKPYYS	KGGEHCYHGH	SIRGVKDSKV	ALSTCNGLHG	MFEDDTFVYM	IEPLELVHDE	240
	KSTGRPHIIQ	KTLAGQYSSQ	MKNLTMRGD	QWFFLSELOH	LKRRKRAVNP	SRGIFPEEMKY	300
	LELMIVNDHX	TYKKRRSSHA	HTNNFAKSVV	NLVDSTYKEQ	LNTRVVLVAV	ETWTEKDQID	360
55	ITTNPVQMLH	EPKSKYRQRI	QHADAHLIS	RVTPHYKRSS	LSYFGGVCSR	TRGVGVNEYG	420
	LFMAVAQVLS	QSLAQNLGIQ	WEPSSRKPCK	DCTESWGGCI	MEETGVSHSR	KPKSKSILEY	480
	RDPLQRGGGA	CLFNRPTKLF	EPTECGNGYV	EAGEECDCGP	HVECYGLCKK	KCSLSNGAHC	540
	SDGPCNINTS	CLFQPRGYEC	RDVNECDIT	EYCTGDSGQC	PPNLEKQDGY	ACNQMQGRCY	600
	NGECKTRDNG	QYIINGTKAA	GSDKFCYEKL	NTEGETKGN	GKDGDRWIQC	SKHDVFCGFL	660
60	LCTNLTRAPR	IGQLQGEIIP	TSFYHQGRVI	DCSGAHVVLD	DDTDVGVVED	GTPOGPMSC	720
	LDRKCLQIQA	LNMSCLPLDS	KGKVCSEHGV	CSNEATCICD	FTWAGTDCSI	RDFVRNLEPP	780
	KDEBKGKPSA	TLNLIIGSIAG	AILVAAIVLG	GTGSGFKNVK	KRRFDPTQQG	PI	

Seq ID NO: 326 DNA sequence
Nucleic Acid Accession #: AK074418.1
Coding sequence: 244-1515

65	1	11	21	31	41	51	
	CTTCTCCAA	GACGGCCGGC	CATGCTCTCC	TCCTCTGCCA	GTCTCTCTCA	CCACTCTCTA	60
70	ACCTGAGAGC	CTGTGGAAAC	TGCCCGTCTC	CCCTCTCTCA	TCAGACACAC	CTGCCTAGGA	120
	AACAGATGGA	AAAAGTGAGG	GACCGGTGAG	TGACTTGTCT	CTAAAGTTTA	TACCAGATGC	180
	AAATGACAGA	GCTGTGAGTT	TGCTGTGCTC	GGAAAGGACC	TGGAAGTCTT	TCTAAGGAGA	240
	GTATGCGGCT	ATTACAGGTA	GCCTTCAGTG	GAGACCTCCA	TCATCAAGTT	CAAGACAGAG	300
	GACTTTACCA	CCTTGGCGGA	TCACTGCCTG	AGCATGGGCC	GGACGTTTAA	GGATGAGACA	360
75	TTCCCGCAG	CAGATTCTTC	CATAGGCCAG	AAGCTGTCTC	AGGAAAAACG	CCTCTCCAAT	420
	GTGATATGGA	AGCGGCCACA	GGATCTACCA	GGGGTCTCTC	CTCACTTCAT	CCTGATGATG	480
	ATAAGCAGAT	TTGACATCCA	ACAAGGAGGC	GCAGCTGACT	GCTGTTCTCT	GGCAGCACTG	540
	GGATCCTTGA	CTCAGAACCC	ACAGTACAGG	CAGAAGATCC	TGATGGTCCA	AAGCTTTTCA	600
80	CACCATATG	CTGGCAITTT	CGTGTTCGGG	TTCTGGCAAT	GTGGCCAGTG	GGTGAAGTGG	660
	GTGATTGATG	ACCGCTTACC	TGTCCAGGGA	GATAAATGCC	TCTTTGTGGG	TCTCGCCAC	720
	CAAAACCAAG	AGTTCTGGCC	CTGCCGTGCT	GAGAAGGCCT	ATGCCAAGCT	GCTCGATGCC	780
	TATTCGATC	TGCACTATGG	CTTCTCTGAG	GATGCCCTGG	TGACCTCAC	AGGAGGCGTG	840
	ATCACCAACA	TCCATCTGCA	CTCTTCCCTT	GTGGACCTGG	TGAAGGCAGT	GAAGACAGCG	900
85	ACCAAGGCAG	GCTCCTGTAT	AACCTGTGCC	ACTCCAAGTG	GGCCACAGAG	TACAGCACAG	960
	GCGATGGAGA	ATGGGCTGGT	GAGTCTCCAT	GCCTACACTG	TGACTGGGGC	TGAGCAGATT	1020
	CAATACGAAA	GGGGCTGGGA	AGAAATTATC	TCCTGTGGGA	ACCCCTGGGG	CTGGGGCGAG	1080
	ACGAATGGA	GAGGCGCTGT	GAGTGTGGG	TCTCAGGAGT	GGGAGGAAAC	CTGTGATCCG	1140

5	CGGAAAGGCC AGCTACATGA GAAACGGGAA GATGGCGAGT TTGGGATGTC GTGTCAAGAT 1200
	TTCCACACAGA AATTCATCGC CATGTTTATA TGTAGCGAAA TTCCAATTAC CCTGGACCAT 1260
	GGAAACACAC TCCACGAAGG ATGGTCCCAA ATAATGTTTA GGAAGCAAGT GATTCTAGGA 1320
	AACACTGCAG GAGGACCTCG GAATGATGCT CAATTCAACT TCTCTGTGCA AGAGCCAAATG 1380
	GAAGGCACCA ATGTTGTCTG GTGGGTCACT GTTGTCTGCA CACCATCAAA TTTGAAAGCA 1440
	GAAGATGCAA AATTTCCACT CGATTTCCAA GTGATTCTGG CTGGCTCACA GAAACACTGT 1500
	CCAAAGCTCA AATAATAAAT TCCGCCGCAA CTTCACCATG ACTTACCATC TGAGCCCTGG 1560
	GAATATGTTT GTGGTTCGAC AGACACGGAG AAAATCAGCG GAGTTCCTGC TCCGAATCTT 1620
10	CCTGAAAATG CCAGACAGTG ACAGGCACCT GAGCAGCCAT TTCAACCTCA GAATGAAGGG 1680
	AAGCCCTTCA GAACATGGCT CCCAACAAAG CATTTTCAAC AGATATGCTC AGCAGGTATG 1740
	GTACCTAGCA CCCAGGGGCC TTACGTGGGA TTGGAGAAAG GGGACCTGAG GGAGGGACAG 1800
	CCCTCAGAGG CCCTTACTGG GATGCAGAGA GGAGAAGTGA CTGTATGGAC TATTTTACCT 1860
	GCCTCTCTTC CTGGATGCTC TCCAGAACTG CTGTGGCTGC CAAGCTCGGT AGAGACGTGG 1920
15	CGCCCCACCC AGTCTCATCC GGGGGACTTC AAGCTGGAAT GCAGAGCTTA GAAAGGGAGG 1980
	GGATAATTAT GGGGTGTGAG GTGCATTGCC CTCTAAATCT TTAACAAGC AATTGGCAGT 2040
	ACCCCGTGAA ACCTTTCTCT CTCTACTCG GCCACCTCC ACCAACCTGG CATCGTTCCT 2100
	CCCGGAGCT AGCCAGCTTC AGAAAGCACA TACAGCATCC TTGCTGCCAA ACCACCTATG 2160
	TGCACACAGG ATTTCTCTAA TGGCTTAATA AACTGTATA AAGAACTCT TGACTTGTCA 2220
20	GAATAAATA GCTGCCAGGG GCTCTGCACA ATGAGCCTCT TACCGTTAAA AAAAAAAAAA 2280
	AAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 327 Protein sequence:
Protein Accession #: BAB85075.1

25	1	11	21	31	41	51	
	MAYYQEPSVE	TSIIKPKDQD	FTTLRDHCLS	MGRTFKDETF	PAADSSIGQR	LLQEKRLSNV	60
	IKWRQDLPG	GPPIFIILDD	SRFDIQQGGA	ADCFWFLAALG	SLTQNPQYRQ	KILMVQSFSS	120
	OYAGIFFRFE	WQCGQWVEVV	IDRLRPVQGD	KCLFVRPRHQ	NQEFWPCLE	KAYAKLGSY	180
30	SDLHYGLEED	ALVBLTGGVI	TNIHLHSPV	DLVKAVKTAT	KAGSLITCAT	PSGPTDTAQA	240
	MENGLVSLHA	YTVTGABQIQ	YRRGWEEIIS	LWNPWNGET	ENRGRWSGDS	QEWETCDPR	300
	KSQHLKRED	GEFWMSCQDF	QQKFIAMPIC	SEIPITLDHG	NTLHEGWSQI	MFRKQVILSN	360
	TAGGPRNDAQ	FNFSVQEPME	GTNVVVCVTV	AVTPSNLKAE	DARFPLDFQV	ILAGSQKHCP	420
35	KLK						

Seq ID NO: 328 DNA sequence
Nucleic Acid Accession #: BC017490.1
Coding sequence: 74-2788

40	1	11	21	31	41	51	
	GTGGGTACAG	TGAACCACTT	TTCCGCGGAA	ACCTGGTGTG	TGCTGTAGTG	GCGGAGAGGA	60
	TCGTGGTACT	GCTATGGCGG	AATCATCGGA	ATCCTTCACC	ATGGCATCCA	GCCCCGCCCA	120
45	GCGTCGGCGA	GGCAATGATC	CTCTCACTTC	CAGCCCTGGC	CGAAGCTCCC	GGCGTACTGA	180
	TGCCCTCACC	TCCAGCCCTG	GCCGTGACCT	TCCACCATTT	GAGGATGAGT	CCGAGGGGCT	240
	CCTAGGCACA	GAGGGGCCCC	TGGAGGAAGA	AGAGGATGGA	GAGGAGCTCA	TTGGAGATGG	300
	CATGGAAAGG	GACTACCGCG	CCATCCCAGA	GCTGGACGCC	TATGAGGCCG	AGGGACTGGC	360
	TCTGGATGAT	GAGGACGTAG	AGGAGCTGAC	GCCGAGTCAG	AGGGAGGCAG	CAGAGCGGGC	420
50	CATGCGGCAG	CGTGACCGGG	AGGCTGGCCG	GGGCTGGGGC	CGCATGCGCC	GTGGGCTCCT	480
	GTATGACAGC	GATGAGGAGG	ACGAGGAGCG	CCCTGCCCGC	AAGCGCCGCC	AGGTGGAGCG	540
	GGCCACGGAG	GAGCGCGAGG	AGGACGAGGA	GATGATCGAG	AGCATCGAGA	ACCTGGAGGA	600
	TCTCAAAGCG	CACCTGTGTC	GCGAGTGGGT	GAGCATGGCG	GGCCCCGGCG	TGGAGATCCA	660
	CCACCGCTTC	AAGAACTTCC	TGCGCACTCA	CGTCGACAGC	CACGCCCCCA	ACGCTCTCAA	720
55	GGAGCGCATC	AGCGACATGT	GCAAGAGAGA	CCGTGAGAGC	CTGGTGTGTA	ACTATGASGA	780
	CTTGCGAGCC	AGGGAGCAGC	TGCTGGCCTA	CTTCTGCCT	GAGGCACCGG	CGGAGCTGCT	840
	GCAGATCTTT	GATGAGGCTG	CCCTGGAGGT	GGTACTGGCC	ATGTACCCCA	AGTACGACCG	900
	CATCACCAAC	CACATCCATG	TCCGATCTTC	CCACTGGCCT	CTGGTGGAGG	AGCTGCGCTC	960
	GCTGAGGCAG	CTGACCTCGA	ACCAGCTGAT	CCGCAACAGT	GGGTGTGTGA	CCAGCTGCAC	1020
60	TGGGCTCCTG	CCCCAGCTCA	GCATGGTCAA	GTACAACATG	AACAAGTGCA	ATTTCTGTCT	1080
	GGGTCTCTTC	TTCGAGTCCC	ASAAACAGGA	GGTGAAACCA	GGCTCTGTCT	CTGAGTGCCA	1140
	GTGCGCCCGC	CCCTTTGAGG	TCAACATGGA	GGAGACCATC	TATCAGAACT	ACCAGCGTAT	1200
	CCGAATCCAG	GAGAGTCCAG	GCAAGATGGC	GGCTGGCCGG	CTGCCCCGCT	CCAAGGACGC	1260
	CATTCTCCTC	GCAGATCTGG	TGGACAGCTG	CAAGCCAGGA	GACGAGATAG	AGCTGACTGG	1320
65	CATCTATCAC	AACAACATAT	ATGGCTCCCT	CAACACTGCC	AATGGCTTCC	CTGTCTTTGC	1380
	CACGTCTCAT	CTAGCCAAAC	ACGTGGCCAA	GAAGGACAAC	AAGGTGCTG	TAGGGGAACT	1440
	GACGAGTGAA	GATGTGAAGA	TGATCACTAG	CCTCTCCAAG	GATCAGCAGA	TGGAGAGAGAA	1500
	GATCTTTGCC	AGCATTTGCT	CTTCCATCTA	TGGTCATGAA	GACATCAAGA	GAGGCTCTGC	1560
	TCTGGCCCTG	TTGCGAGGGG	AGCCCCAAAA	CCCAGGTGGC	AAGCACAAAG	TACGTGGTGA	1620
70	TATCAACGTG	CTCTGTGTGG	GAGACCTTGG	CACAGCGAAG	TGCGAGTTTC	TCAAGTATAT	1680
	TGAGAAAGTG	TCCAGCCGAG	CCATCTTCAC	CACCTGGCCG	GGGGCGTCCG	CTGTGGGCTC	1740
	CACGGCGTAT	GTCACGCGGC	ACCCTGTCTG	CAGGGAGTGG	ACCTTGGAGG	CTGGGGCCCT	1800
	GGTTCTGGCT	GACCGAGGAG	TGTGTCTCAT	TGATGAATTT	GACAAGATGA	ATGACCAGGA	1860
	CAGAACCAGC	ATCCATGAGG	CCATGGAGCA	ACAGAGCATC	TCCATCTCGA	AGGCTGGCAT	1920
75	CGTCACCTCC	CTGACGGCTC	GCTGCAAGGT	CATTGTCTGC	GCCCAACCCA	TAGGAGGGGC	1980
	CTACGACCCC	TCGCTGACTT	TCTCTGAGAA	CGTGGACCTC	ACAGAGCCCA	TCATCTCAGG	2040
	CTTTGACATC	CTGTGTGTGG	TGAGGGACAC	CGTGGACCCA	GTCCAGGACG	AGATGCTTGG	2100
	CCGCTTCTGT	GTGGGCGAGC	AOGTGAGACA	CCACCCAGCG	AACAAGGAGG	AGGAGGGGCT	2160
	GGCCAAATGG	AGCGCTGTCT	AGCCCCCAT	GCCCAACACG	TATGGCTGGG	AGCCCCCTGC	2220
80	CCAGGAGGTC	CTGAAGAAGT	ACATCATCTA	CGCCAAGGAG	AGGGTCCACC	CGAAGCTCAA	2280
	CCAGATGGAC	CAGGACAAGG	TGGCCAAGAT	GTACAGTGAC	CTGAGGAAAG	AATCTATGGC	2340
	GACAGGCAGC	ATCCCCATTA	CGGTGCGGCA	CATGAGTCC	ATGATCCGCA	TGGCGGAGGC	2400
	CCAGCGCGCG	ATCCATCTGC	GGGACTATGT	GATGGAAGAC	GACGTCAACA	TGGCCATCCG	2460
	CGTGATGCTG	GAGAGCTTCA	TAGACACACA	GAAGTTCAGC	GTCAATGCGA	GCATGCGCAA	2520
85	GACTTTTGCC	CGCTACCTTT	CATTCCGGCG	TGACAACAAT	GAGCTGTGTC	TCTTCTACT	2580
	GAAGCAGTTA	GTGGCAGAGC	AGGTGACATA	TCAGCGCAAC	CGCTTTGGGG	CCCAGCAGGA	2640
	CACATTTGAG	GTCCCTTGAGA	AGGACTTGGT	GGATAAGGCT	CGTCAGATCA	ACATCCACAA	2700

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

CCCTCTGCA TTTTATGACA GTGAGCTCTT CAGGATGAAC AAGTTCAGCC AGCACTGAA 2760
AAGGAAAATG ATCTCTGAGC AGTCTCTGAGG CCTATGATGCA TCCTAAGGA TTCCTTGGGA 2820
TTCTGGTTTG GGGTGGTTCAG TGCCCTCTGT GCTTTATGGA CAAAAACCA GAGCACTTGA 2880
TGAACTGGG GTACTAGGGT CAGGGCTTAT AGCAGGATGT CTGGCTGCAC CTGGCATGAC 2940
TGTITGTTTC TCCAGAGCTG CTTTGTGCTT CTCACCTTTG GGTGGGATGC CTGGCAGTGT 3000
TGCTTACTT GGTGTCTGAA CATTTGCA CCTCGAGTG CTTTGTCTCC ACTCAGTACC 3060
TTGGATCAGA GCTGCTGAGT TCAGGATGCC TGGGTGAGT TTAGGTGTTA GCCTTCTTAC 3120
ATGGATGTCA GGAGAGCTGC TGCCCTCTTG CGGTGAGTTG CGTATTCAGG CTGCTTTTGC 3180
TGCCCTTGGC CAGAGAGCTG GTTGAAGATG TTTGTAATCG TTTTCAGTCT CTGCAGGTT 3240
TCTGTGCCCC TGTGTGGAAG GAGGGCAGCA CAGTGCCAGC GCAGGCTTCT GGGCTCCTCA 3300
GTGCGAGGGG TGGGATGTGA GTCATGCGGA TTATCCACTC GCCACAGTTA TCAGCTGCCA 3360
TTGCTCCCTG TCTGTTTCCC CACTCTCTTA TTTGTGCAAT CGGTTTGGTT TCTGTAGTTT 3420
TAATTTTAA TAAAGTTGAA TAAATATAA AAAAAA AAAAAA

Seq ID NO: 329 Protein sequence:
Protein Accession #: AAH17490.1

1 11 21 31 41 51
MAESSESEFTM ASSPAQRRRG NDPLTSSPGR SSRRTDALTS SPGRDLPPFE DESEGLLGTE 60
GPLEEEDDGE ELIGDGMERD YRAIPELDAY EAEGALDDE DVEELTASQR EAAERAMRQR 120
DREAGRGLGR MRRGLLYDSO EDEERPARK RRQVERATED GEEDEEMIES IENLEDLKGH 180
SVREWVSMAG PRLEIHRHFK NFLRTHVDSH GNVFKERIS DMCKENRESL VVNYEDLAAR 240
EHLVAYELPE APAELLQIFD EAALVVLAM YPKYDRITNH IHRVISHLPL VEELRLSLRL 300
HLNQLITSG VVTSGTGLVP QLSMVKYNCN KCNFLVGPFC QSQNQEVKPG SCPECCSAGP 360
FEVMEETIY QNYQRIHQE SPKVAAGRL PRSKDAILLA DLVDSCKPGD EIELTGIYHN 420
NYDGLNLTAN GFPVPATVIL AMHVAKDNK VAVGELTDED VMITSLSKD QQIGKIFAS 480
IAPSIYGHED IKRGLALALP GGEKPNPGGK HKVRGDNVL LCGDPGTAKS QFLKYIEKVS 540
SRAIFTTGGG ASAVGLTAYV QRHPVSEWT LEAGALVLAD RGVCLIDEFD KMNDQRTSI 600
HEAMEQSSIS ISKAGIVTSL QARCTVIAAA NPICGRYDPS LTFSENVDLT EPIISRFDIL 660
CVVRDVTDPV QDEMLARFVQ GSHVRHHPNS KEEGLANGS AAEPMAMNTY GVEPLPQEV 720
KKYIYAKER VHPKLNQMDQ DKVAKMYSYL RKESMATGSI PITVRHIESM IRMAEAHARI 780
HLRDYVIEDD VMAIRVMLE SFIDTQKFSV MRSMRKTFAR YLSFRDNNE LLLFILKQLV 840
AEQVYQQRNR PGAQDITIEV PEKDLVDKAR QINIHLNLSAP YDSELFMRNK FSHDLKRMI 900
LQQF

Seq ID NO: 330 DNA sequence
Nucleic Acid Accession #: M17254
Coding sequence: 257-1645

1 11 21 31 41 51
GTCGCGCGGT GTCCGCGCCC GCGTGTGCCA GCGGCGGTGC CTGCGCGTG GCGGCGGAGC 60
CGGGTGCAC TAACCTCCCTC GGCGCGGACG GCGGCGCTAA CCTCTCGGTT ATTCAGGAT 120
CTTTGGAGAC CCGAGGAAAG CCGTGTGAC CAAAGCAAG ACAATGACT CACAGAGAA 180
AAGATGGCA GAACCAAGG CACTAAAGC CGTCAGTTC TGAACAGCTG GTAGATGGGC 240
TGGCTTACTG AAGGACATGA TTCAGACTGT CCGGACCCA GCAGCTCATA TCAAGGAAGC 300
CTTATCAGTT GTGAGTGAAG ACCAGTCTGT GTTTGAGTGT GCCTACGGAA CGCCACACCT 360
GGCTAAGACA GAGATGACCG GTCTCTCTC CAGCGACTAT GGACAGACTT CCAAGATGAG 420
CCCAAGGTC CCTCAGCAGG ATTGGCTGTC TCAACCCCA GCCAGGATCA CCATCAAAAT 480
GGAATGATC CTAGCCAGG TGAATGGCTC AAGGAACCTC CCGTATGAAT GCAGTGTGGC 540
CAAAGCGGG AAGATGGTGG GCAGCCAGA CACCGTTGGG ATGAACACG GCAGCTACAT 600
GGAGGAGAAG CACATGCCAC CCCCAACAT GACCACGAAC GAGCGCAGAG TTATCGTGCC 660
AGCAGATCCT ACAGCTATGA GTACAGACCA TGTGCGGAG TGGCTGGAGT GGGCGGTGAA 720
AGAATATGGC CTTCAGACG TCAACATCTT GTTATTCCAG AACATCGAT GGAAGGAAC 780
GTCAAGATG ACCAAGGAGC ACTTCCAGAG GCTCACCCCA AGCTACAACG CCGACATCCT 840
TCTCTCATAT CTCCACTACC TCAGAGAGAC TCCTCTTCCA CATTTGACTT CAGATGATGT 900
TGATAAAGCC TTACAAACAT CTCCAGGTTT AATGCATGCT AGAAACACAG ATTTACCATA 960
TGAGCCCCC AGGAGTACG CCGGACCGG TCACGCCAC CCCACGCCCC AGTCGAAAGC 1020
TGCTCAACCA TCTCTCTCCA CAGTGCCCAA AACTGAAGAC CAGGTCCTC AGTTAGATCC 1080
TTATCAGATT CTTGACCAA CAAGTAGCCG CTTGCAAAAT CCAGGCAGTG CCAGATCCA 1140
GCTTTGSCAG TTCTCTCTGG AGCTCTGTG GGACAGCTCC AACTCCAGCT GCATCACCTG 1200
GGAAGGCACC AACGGGGAGT TCAAGATGAC GGATCCCGAC GAGGTGGCCC GGGCTGGGG 1260
AGAGCGGAAG AGCAAAACCA ACATGAACAT CGATAAGCTC AGCGCGGCC TCCGTTACTA 1320
CTATGACAG AACATCATGA CCAAGGTCCA TGGGAAGCGC TACGCTACA AGTTGCACTT 1380
CCACGGGATC CCGCAGGCC TCCAGCCCA CCCCCGGAG TCACTCTGT ACAAGTACC 1440
CTCAGACCTC CGTACATGG GCTCCTATCA CGCCCAACCA CAGAAGATGA ACTTTGTGGC 1500
GCCCCACCTT CCAGCCCTCC CCGTGACATC TTCCAGTTT TTTGCTGCC CAAACCCATA 1560
CTGGAATTCA CCAACTGGGG GTATATACCC CAACACTAGG CTCCCCACCA CCATATGCC 1620
TTCTCATCTG GGCACCTACT ACTAAGACC TGGCGGAGGC TTTTCCCATC AGCGTGCAAT 1680
CACAGCCCA TCGCCAGAA CTCTATCGGA GAACATGAAT CAAAAGTGCC TCAAGAGGAA 1740
TGAAAAAAGC TTTACTGGGG CTGGGGAAGG AAGCGGGGA AGAGATCCA AGACTCTGG 1800
GAGGGAGTTA CTGAAGTCTT ACTACAGAAA TGAGGAGGAT GCTAAAAATG TCACGAATAT 1860
GGACATATCA TCTGTGGACT GACCTTGTA AAGACAGTGT ATGTAGAAGC ATGAAGTCTT 1920
AAGGACAAG TGCCAAAGAA AGTGGTCTTA AGAAATGTAT AAACCTTAGA GTAGAGTTTG 1980
AATCCCAAT ATGCAAACTG GGATGAACT AAAGCAATAG AAACAACCA GTTTGACCT 2040
AACATACCGT TTATAATGCC ATTTTAAAG AAACCTACCT TATTTAAAA TAGTTTCATA 2100
TCAAAAAACA GAGAAAGAG ACGAGAGAGA CTGTGCCCCA TCAACAGACG TTGATATGCA 2160
ACTGATGGC ATGTGCTGTT TTTGTTGAAA TCAAAATCAT TCGTTTGAT GGACAGCTGT 2220
CAGCTTTCTC AAACGTGAA GATGACCCAA AGTTTCCAA TCCTTTACAG TATTACCGGG 2280
ACTATGAAC TAAAGGTGGG ACTGAGATG TGTATAGAGT GAGCGTGTGA TTGTAGACAG 2340
AGGGGTGAAG AAGGAGGAGG AAGAGGCAGA GAAGGAGGAG ACCAGGCTGG GAAAGAACT 2400
TCTCAAGCAA TGAAGACTGG ACTCAGGACA TTTGGGGACT GTGTACAATG AGTTATGGAG 2460
ACTCGAGGGT TCACTGAGTC AGTGTATAT CAAACCCAGT GTTAGGAGAA AGGACACAGC 2520
GTAATGAGGA AAGGGAAGTA GTAGAATTCA GAAACAAAAA TGCGCATCTC TTTCTTTGTT 2580
TGTCAAATGA AAATTTTAA TCGAATTGTC TGATATTTAA GAGAAACATT CAGGACCTCA 2640
TCATTATGTG GGGGCTTTGT TCTCCACAGG GTCAGGTAAG AGATGGCCTT CTTGGCTGCC 2700

ACAATCAGAA ATCAOCCAGG CATTTTGGGT AGGCGGCTC CAGTTTTCCT TTGAGTCGGG 2760
 AACGCTGTGC GTTTGTGAGA ATGAAGTATA CAAGTCAATG TTTTTCOCCT TTTTATATA 2820
 ATAATTATAT AACTTATGCA TTTATACACT ACGAGTTGAT CTGGCCAGC CAAGACACA 2880
 CGACAAAAGA GACAATCGAT ATAATGTGGC CTGGAATTTT AACTCTGTAT GCCTAATGTT 2940
 TACAATATGA AGTTATTAGT TCTTAGAATG CAGAATGTAT GTAATAAAT AAGCTTGGCC 3000
 TAGCATGGCA AATCAGATT ATACAGGAGT CTGCATTTGC ACTTTTTTTA GTGACTAAAG 3060
 TTGCTTAATG AAAACATGTG CTGAATGTTG TGGATTTTGT GTTATAATT ACTTTGTCCA 3120
 GGAACCTTGT CAAGGAGAG CCAAGGAAAT AGGATGTTG GCACCC

Seq ID NO: 331 Protein sequence
 Protein Accession #: AAA52398

1 11 21 31 41 51
 15 MIQTVDPDAA HIKEALSUVS EDQSLFECAY GTPHLAKTEM TASSSSDYGQ TSKMSPRVPO 60
 QDHLSPPAR VTIKMCNPS QVNGSRNSPD ECSVAKGGKM VGSFDTVGMN YGSYMSEKHM 120
 PPNMTTNER RVIVPADPTL WSTDHVRQWL ENAVKEYGLP DVNILLFQNI DGKELCKMTK 180
 DDFQLRLTSPY NADILLSHLH YLRETPLPHL TSDDVDKALQ NSPLMHARN TDLFVFPFR 240
 20 SAWTGHGHT POKAAQPSF STVPKTEDOR POLDPYQILG PTSSRLANPG SGQIQWLQFL 300
 LELLSDSNS SCITWGTNG EFKMTDPDEV ARRWGERKSK PNMNYDKLSR ALRYYDKNI 360
 MTKVHGKRYA YKDFPHGIAQ ALQPHFPSS LYKYPDLFY MGSYHAHPQ MNFVAPHPA 420
 LPVTSSSFFA APNPYNSET GGIYPNTRLP TSHMPSHLGT YY 462

Seq ID NO: 332 DNA sequence
 Nucleic Acid Accession #: NM_000020
 Coding sequence: 283-1794

1 11 21 31 41 51
 30 AGGAAACGGT TTATTAGGAG GGAGTGGTGG AGCTGGGCCA GGCAGGAAGA CGCTGGAATA 60
 AGAAACATTT TTGCTCCAGC CCCCATCCCA GTCCCGGGAG GCTGCGCGC CAGCTGCGCC 120
 GAGCGAGCCC CTCCCGGGT CCAGCCCGGT CCGGGGCGCG GCCGAGCCCC AGCCCGCGT 180
 CCAGCGCTGG CGGTGCAACT GCGGCCGCGC GGTGGAGGGG AGGTGCCCC GGTCCGCGA 240
 35 AGGCTAGCGC CCGGCCACCC GCAGAGCGGG CCCAGAGGGA CCATGACCTT GGGCTCCCCC 300
 AGGAAAGGCC TTCTGATGCT GCTGATGGCC TTGTTGACCC AGGAGAGCCC TGTGAAGCG 360
 TCTCGGGGCC CGCTGGTGAC CTGCACTGTG GAGAGCCAC ATTGCAAGGG GCCTACCTGC 420
 CGGGGGGCTT GGTGCACAGT AGTGTGTGTG CCGGAGGAGG GGAGGCACCC CCAGGAACAT 480
 CGGGGCTGCG GGAACCTGCA CAGGAGCTC TGCAAGGGGC GCCCACCGA GTTCGTCAAC 540
 40 CACTACTGCT GCGACAGCCA CCTCTGCAAC CACAAGCTGT CCCTGTGTCT GAGGCCACC 600
 CAACCTCCTT CCGAGCAGCC GGGAAAGAT GGCACAGTGG CCCTGATCCT GGGCCCGCTG 660
 CTGGCCTTGG TGCCCTGTGT GGCCTTGGGT GTCTTGGGCC TGTGGCATGT GCGACGGAGG 720
 CAGGAGAGCG AGCGTGGCT GCACAGCGAG CTGGGAGAGT CCAGTCTCAT CCTGAAAGCA 780
 TCTGAGCAGG GCGACACGAT GTTGGGGGAC CTCTGAGCA GTGACTGCAC CACAGGAGT 840
 45 GGCTCAGGGC TCCCTTCTCT GGTGCAGAGG ACAGTGGCAC GGCAGGTTGC TTTGGTGGAG 900
 TGTGTGGGAA AAGGCGCTTA TGGCGAAGTG TGGCGGGGCT TGTGGCACGG TGAGAGTGTG 960
 GCGCTCAAGA CTCTCTCCTC GAGGGATGAA CAGTCTGTGT TCCGGGAGAG TGAGATCTAT 1020
 AACACAGTAT TGCTCAGACA GCACCAATC CTAGGCTTCA TCGCCTCAGA CATGAOCTCC 1080
 CGCAACTCGA GCACGAGCT GTGGCTCATC ACGCACTACC ACGAGCACGG CTCCTCTAC 1140
 50 GACTTTCTGC AGAGACAGAC GCTGGAGCCC CATCTGGCTC TGAGGCTAGC TGTGTCCGCG 1200
 GCATGCGGCG TGGCGCACTT GCACGTGGAG ATCTTGGGTA CACAGGGCAA ACCAGCCATT 1260
 GCCCACCGCG ACTTCAAGAG CCGCAATGTG CTGGTCAAGA GCAACCTGCA GTGTTGCATC 1320
 GCGGACCTGG GCCTGGCTGT GATGCACTCA CAGGGCAGCG ATTACCTGGA CATCGGCAAC 1380
 AATCCGAGAG TGGGCACCAA GCGGTACATG GCACCGAGG TGCTGGAAGA GCAGATCCCG 1440
 55 ACGGACTGCT TTGAGTCTTA CAAGTGGACT GACATCTGGC CCTTTGGCCT GGTGCTGTGG 1500
 GAGATTGCCC GCGCGACCAT GGTGAATGGC ATCTGTGGAG ACTATAGACC ACCCTTCTAT 1560
 GATGTGGTGG CCAATGACCC CAGCTTTGAG GACATGAAGA AGGTGGTGTG TGTGGATCAG 1620
 CAGACCCCCA CCATCCCTAA CCGGCTGGCT GCAGACCGGG TCCTCTCAGG CCTAGCTCAG 1680
 ATGATGCGGG AGTGTGTTGA CCAAAACCCC TCTGCCGAC TCACCGCGCT GGGATCAAG 1740
 60 AAGACACTAC AAAAAATTAG CAACAGTCCA GAGAAGCCTA AAGTGATTCA ATAGCCACAG 1800
 AGCACTGAT TCCTTTCTGC CTGCGGGGGG CTGGGGGGGT GGGGGGCGAT GGATGGTGCC 1860
 CTATCTGGGA AGAGGTAGTG TGAGTGTGGT GTGTGCTGGG GATGGGCGAG TGCGCCTGCC 1920
 TGCTCGGCCG CAGGCCACCC CAGCCAAAAA TACAGCTGGG CTGAAACCTG ATCCCTGCT 1980
 GTCTGGCCTG CTCAAAGCGG CAGGCTCCCT GACGCTGGC TCTCTCCCA CCCCTATGGC 2040
 65 CAGCATGGTG CACCCCTTAC CACTCCCGGG ACAGGATGCA AAGAGGCTC CAGAGTCAGA 2100
 GTGCCAAGCG AGGGAATCCC AGTCCAGAG TCAGAGCCCG GGCTGCACT TTGCCCTCTG 2160
 CCCTTGATCA ACCCACTGC CCCACAGAG CTGCCAGGGT GGCACAGGGC CCGTCCAGC 2220
 CCCTGGCACA CACTCCCTG CCAGGCTCA GCTCTAGCA TAAGCTCCAG AGAGCCAGGG 2280
 CCCATCAGTT TCTCTCTGTG GATTGTATC TCAGCTCCAT GATGCTTGG GCTTTCTGTC 2340
 70 TCCTCAACAA GAGTGCAGCT TGCTGAATGT CAGCTGCCTG AGAGAGCTGG GGCCTGACTT 2400
 ACTAGGGCAT TAAATCCTAA GAGGTCTTAC TGAGGTGTGG CAGGATCACA GGCCAGTGG 2460
 AAAAGGGCAG GTCAGATGGG CAAGGCCAG GACTTTCAGA TTAAGTGA GAATATCGAG 2520
 GCCAAGCATG GCAGGGGAA GGTCAATGGG TGTCAAGAGA CCCAGGCTG ACCCGGATG 2580
 TTTGCTCCAT GTGACAAAG CAGGCTGTC TCAGGACCTT TCTTTTCTT TTTTCTTCT 2640
 75 TTTTCTTTT GACACGGAGT TTCGCTCTG TTGTCCAGGC TAGAGTGCAA TGCCATGATC 2700
 CCAGCTCACC GCAACGTCTA CCTCCAGGT TCAATCATT CTCTTGCTC AGACTCCGGA 2760
 GTAGCTGGGA TTAAGGCAC ATGCCACCAT GCGTGGCTAA TTTTGTATAT TTAGTAGAAA 2820
 CAGGGTTTCA CCATGCTGCG CATGCTGGT CTGCAACTCC TGACCTCAGG TGTTCACCT 2880
 ACCTCAGCCT CCCAAAGTGC TGGGGTTACA GGTGTGAGCC ATGCGCCTG GCCAGGACCT 2940
 80 TTGTTCTTA TCTACATATT GGAAGATTG GTCTGTATGT CCTTTGAGGC TTTCTTAGCT 3000
 CTAGTTCTCT GACACTTCAG CCTATATCAC AGCTAACTTC YTCAGTCTCA TCTATTCTT 3060
 ATGCTCCAGC CCTGGCAAT TTGCTCAAG ATGGGGGTTT GAAAATAACT TTACCTGACT 3120
 CAAGGAGTGT CTGAGCACC TCCTAGTCTA AGTCTGCAAG CTCAGTCTT TGCTTAAAC 3180
 CATGCCAGTG GCCACCTTGG GGCTCAGACA GCTCTGGGCC TTTTGACCA AGGCCAGCCC 3240
 85 CTGCGCCTCT CTGTGCATA GTCTTCTCTG CCCAGGACT GCGGGGCGG TTCTCCAG 3300
 GCTTCCAGG CTCAAAAGAA ATTGGCTCC ATCCAGAAG GCTCCAGCTC CCTTCTGGC 3360
 CCTGGCTTC AGGCCACAC CCTGGGCCA GGCACAGAGA GTGTGTCTCA GGAAATTTCA 3420
 ATGGGCTCTA GAGAGACACA CAGAAAGTTT GGGCATTTGG GAAATTTTCA AGRTGTATG 3480

5 TATGGYTAC GTATGGWGCA GGTGTGCTG GTCCYKGGGT GCAGGGAAGT GGGCTGCAGG 3540
 GAAGTGGATT GGAGGGGAGC TTGAGGAATA TAAGGAGCGG GGGTGGAGAC TCAGGCTATG 3600
 GACAAGGACA GCCCAAGGT TGGGAAGACC TGGCCTTAGT COTCTCAGC CTAGGGCAGG 3660
 GCAGTGAAGA AAGCTCTCCC CGCTCCTGCT GTAATGACCC AGAGTAGCCT CCCAGGCGG 3720
 GCATCTTATG TGTGTCTTCC ACCATCTCTCA TGGTGGCACT TTTCTAGGCC TGTCTCCAG 3780
 CATTGTGCAA GGCTGGGAAG AGAACCCAGGA AGTGAACCTG GGTGAAACA GAAAGCTCAA 3840
 TGGATGGGCT AGGTCCAG ATCATTAGGG CAGAGTTTGC ACGTCTCTG GTTCACTGGG 3900
 AATCCACCCA GCCCAGCAAT CATCTCCTC TTTGAAGGAT TTTWATTCT ACTGGGTTTT 3960
 GGAACAACT CCTGCTGAGA CCCACAGCC AGAACTGAA AGCAGCAGCT CCCCAAGGCC 4020
 10 TGGAAATCC CTAAGAGAAG GCCTGGGGGA MAGGAATGG AGTGACAGG GACAGGTAGA 4080
 GAGAAGGGG CCCAATGGCC AGGGAGTGAA GGAGGTGGG TTGCTGAGAG CAGTCTGCAC 4140
 ATGCTTCTGT CTGAGTGCAG GAAAGTGTTC CAGGCTCGAA ATTACACTTC TCGTACCTGG 4200
 AGACGCTGTT TGTGGGAGCA CTGGGCTCAT GCCTGGCACA CAATAGGTCT GCAATAAAC 4260
 15 ATGGTTAAAT CCTGAAAAA AAAAAAAA

Seq ID NO: 333 Protein sequence
 Protein Accession #: NP_000011

20 1 11 21 31 41 51
 MTLGSPKGL LMLLMALVTQ GDPVKPSRGP LVTCTCESPH CRGPTCRGAW CTVVLVREEG 60
 RHPQERLRCGR NLHRELRCGR PTEFVNHYCC DSHLCNHNVS LVLEATQPPS EQPGTDGQLA 120
 25 LILGVLALL ALVALGVLGL WHVRRRQEKQ RGLHSELGES SLILKASEQG DTMLGDLIDS 180
 DCTTGGSSGL PFLVQRTVAR QVALVECVGK GRYGEVWRGL WHGESVAVKI FSSRDEQSWF 240
 RETEIYNTVL LRHNDILGFI ASDMTSRNSS TQLWLITHYH EHGSYLDILO RQTLPHLAL 300
 RLAVSAACGL AHLHVEIFGT QGKPAIAHRD FKSRNVLVKS NLQCCIALDLG LAVMHSQSGD 360
 YLDIGNNPRV GTIKRYMAPEV LDEQIRTDCE ESKYKWDIWA FGLVLWEIAR RTIVNGIVED 420
 YRPPFYDVP NDPSFEDMKK VVCVDQQTPT IPNRLAADPV LSLAQMMRE CWPYNPSSARL 480
 30 TALRIKKTLO KISNSPEKPK VIQ

Seq ID NO: 334 DNA sequence
 Nucleic Acid Accession #: NM_004126.1
 Coding sequence: 108-329

35 1 11 21 31 41 51
 GCACGAGCT CGTGCCGGCC TTCAGTTGTT TCGGGACGCG CCGAGCTTCG CCGCTCTTCC 60
 AGCGGCTCCG CTGCAGAGC TAGCCCGAGC CCGGTTCTGG GCGGAAATG CCGCCCTTC 120
 40 ACATCGAAGA TTTGCCAGAG AAGGAAAAAC TGAAATGGA AGTTGAGCAG CTTCGCAAG 180
 AAGTGAAGTT GCAGAGACAA CAAGTGTCTA AATGTTCTGA AGAATAAAG AACTATATTG 240
 AAGAACGTTT TGGAGAGGAT CCTCTAGTAA AGGGAATTCC AGAAGACAAG AACCCCTTTA 300
 AAGAAAGAG CAGCTGTGTT ATTTCAATAA TAACITGGGA GAAACTGCAT CTTAAGTGA 360
 45 AGAAGTAGTT TGTTTTAGTT TTCCAGATA AAACCAACAT GCTTTTAAAG GAAGGAAGAA 420
 TGAATATAAA AGGAGACTTT CTTAAGCACC ATATAGATAG GGTATGTAT AAAAGCATAT 480
 GTGCTACTCA TCTTGTCTCA CTATGCAGTC TTTTAAAGA GAGCAGAGAG TATCAGATGT 540
 ACAATTATGG AAATAAGAAC ATTACTTGAG CATGACACT CTTTCAGTAT ATTGCTTGAT 600
 GCTTCAATA AAGTTTGTCT TT

50 Seq ID NO: 335 Protein sequence
 Protein Accession #: NP_004117.1

55 1 11 21 31 41 51
 MPALHIEDLP EKEKLKMEVE QLRKEVKLQR QVSKSCSEI KNYIERSGE DPLVKGIPED 60
 KNPFKEKGSC VIS

60 Seq ID NO: 336 DNA sequence
 Nucleic Acid Accession #: NM_005795
 Coding sequence: 555-1940

65 1 11 21 31 41 51
 GCACGAGGGA ACAACCTCTC TCTCTSCAGC AGAGAGTGTC ACCTCTGCT TTAGGACCAT 60
 CAAGCTCTGC TAACTGAATC TCATCCTAAT TGCAGGATCA CATTGCAAG CTTTCACTCT 120
 TTCCACCTT GCTTGTGGGT AAATCTCTTC TGGGAATCT CAGAAAGTAA AGTTCCATCC 180
 TGAGAATATT TCACAAGAA TTTCTTAAG AGCTGGACTG GGTCTTGACC CTTGGAATTT 240
 70 AAGAAATCT TAAGACAAT GTCAAATATG ATCCAAGAGA AAATGTGATT TGAGTCTGGA 300
 GACAAATGTG CATATCGTCT AATAATAAAA ACCCATACTA GCCTATAGAA AACAATATTT 360
 GAATAATAAA AACCCATCAT AGCCTATAGA AAACAATATT TGAAGATTG CTACCACTAA 420
 AAAGAAACT ACTACAACCT GACAAGACTG CTGCAAACTT CAATTGGTCA CCACAACCTG 480
 ACAAGGTGTC TATAAAACAA GATTGCTACA ACTTCTAGTT TATGTTATAC AGCATATTTT 540
 75 ATTTGGGCTT AATGATGGAG AAAAGTGTA CCCTGTATT TCTGTTCTC TTGCCTTTT 600
 TTATGATTCT TGTATACAGA GAATTAGAAG AGAGTCTGGA GGAATCAAT CAGTGGGAG 660
 TTACTAGAAA TAAATCATG ACAGCTCAAT ATGAATGTTA CCAAAAGATT ATGCAAGACC 720
 CCATTCAACA AGCAGAGGCG GTTTACTGCA ACAGAACCTG GGAATGATGG CTCTGCTGGA 780
 80 ACGATGTGTC AGCAGGAAC GAATCAATGC AGCTCTGCCC TGATTACTTT CAGGACTTTG 840
 ATCCATCAGA AAAAGTTACA AAGATCTGTG ACCAAGATGG AAATGGTGT AGACATCCAG 900
 CAAGCAACAG AACATGGACA AATTATACCC AGTGTAAAT TAACACCCAC GAGAAAGTGA 960
 AGACTGCCT AAATTGTTT TACCTGACCA TAATTGGACA CGGATGTCT ATTGCATCAC 1020
 TGCTTATCTC GCTTGGCATA TTCTTTTATT TCAAGAGCCT AAGTGTCCAA AGGATTACCT 1080
 85 TACACAAA TCTGTTCTTC TCAATTTGTT GTAACCTCTG TGAACAATC ATTCACCTCA 1140
 CTGCAGTGGC CAACAACAG GCCTTAGTAG CCACAATCC TGTATGTTC AAAGTGTCCC 1200
 AGTTCAITCA TCTTACCTG ATGGGCTGTA ATTACTTTTG GATGCTCTGT GAAGGCATTT 1260
 ACCTACACAC ACTCATTTGT GTGGCGTGT TTGCAGAGAA GCAACATTTA ATGTGATATT 1320
 ATTTCTTGG CTGGGATTT CCACTGATTC CTGCTGTAT ACATGCCATT GCTAGAAGCT 1380

TATATTACAA TGCAATTCGC TGGATCAGTT CTGATACCCA TCTCCTCTAC ATTATCCATG 1440
 GCCCAATTGG TCGTGTCTTA CTGGTGAATC TTTTTTCTT GTTAAATATT GTACGGGTTC 1500
 TCATCACCAG GTTAAAGGTT ACACACCAAG CGGAATCCAA TCTGTACATG AAAGCTGTGA 1560
 GAGCTACTCT TATCTTGGTG CCATTGCTTG GCAITGAATT TGTGCTGATT CCATGGCGAC 1620
 CTGAAGGAAA GATTTCAGAG GAGGTATATG ACTACATCAT GCACATCCTT ATGCACITCC 1680
 AGGGTCTTTT GGTCTCTACC ATTTTCTGCT TCTTTAATGG AGAGGTTCAG GCAATCTCTA 1740
 GAAGAACTCG GAATCAATAC AAAATCCAAAT TTGGAACAG CTTTCCAAAC TCAGAAGCTC 1800
 TTCGTAGTGC GTCTTACACA GTGTCAACAA TCASTGATGG TCCAGGTATAT AGTCATGACT 1860
 GTCCTAGTGA ACACCTAAAT GGAAAAAGCA TCCATGATAT TGAATATGTT CTCTTAAAC 1920
 CAGAAATTTT ATATAATTGA AAATAGAAGG ATGGTTGTCT CACTGTTTGG TGCTTCTCCT 1980
 AACTCAAGGA CTGGACCCCA TGACTCTGTA GCCAGAAGAC TTCAATATTA AATGACTTTG 2040
 GGGAAATGTC TAAAGAAGAG CCTTCACATG AAATTAGTAG TGTGTTGATA AGAGTGTAA 2100
 ATCCAGCTCT ATGTGGGAAA AAAGAAATCC TGGTTTGTAA TGTGTTGTCAG TAAATCTCC 2160
 CACTATGCCT GATGTGAGCG TACTAACCTG ACATCACCAG GTGTGGAATT GGAGAAAAGC 2220
 ACAATCAACT TTCTGAGCT GGTGTAAGCC AGTTCAGCA CACCATTTGAT GAATTCACAC 2280
 AAATGGCTGT AAAACTAAAC ATACATGTTG GGCATGATTC TACCTTTATT CSCCCCAAGA 2340
 GACCTAGCTA AGGTCTATAA ACATGAAGGG AAAATTAGCT TTTAGTTTAA AAACCTTTA 2400
 TCCCATCTTG ATGTGGGCGC TTGACTTTTT TTTTTCCTCA GAGTCCCGTA GTCCCTTTTG 2460
 TAATACCTCT CTCAAAATGGA CAATACCAGA AGTGAATAT CCGTCTGGC TTTCTTTCT 2520
 CTATGAAAAG CAATCTGAGT CAATTGTTAT GATCTACTCA TTTGCTGACA CATCAGTTAT 2580
 ATCTGTGGC ATATCCATTG TGGAAACTGG ATGAACAGGA TGTATAATAT GCAATCTTAC 2640
 TTCTATATCA TTAGAAAAC ATCTTAGTTG ATGCTACAAA ACACCTGTGC AACCTCTCC 2700
 TGTCTTACCA AACAGTGGGA GGGAAATCCT AGCTGTAAT ATAAATTTTG CCCTTCCATT 2760
 TCTACTGTAT AAACAAATTA GCAATCATTT TATATAAAGA AAATCAATGA AGGATTTCTT 2820
 ATTTTCTTGG AATTCTTGTAA AAAGAAATTT TGAATAATGA GCTGTAAAT ACTCCATTAT 2880
 TTTATTTTAT AGTCTCAAAAT CAAATACATA CAACCTATGT AATTTTAAAG GCAATATAT 2940
 AATGCAACAA TGTGTGTATG TTAATATCTG ATACTGTATC TGGCTGATT TTTTAAATAA 3000
 AATAGAGTCT GAATGCT

Seq ID NO: 337 protein sequence
 Protein Accession #: NP_005786.1

1 11 21 31 41 51
 MEKKCTLYPL VLLPFFMILV TAELEESPED SIQLGVTRNK IMTAQYECYQ KIMQDPIQQA 60
 EGVYCNRTWD GMLCNMDVAA GTESMQLCPD YFQDFDSEK VTKICDQDGN WFRHPASNRT 120
 WNYTQCNVN THEKVKTALN LFYLTIIHGG LSLASLLISL GIFFYFKSL SQRITLHKNL 180
 FFSFVNCNVV TIHLTAVAN NQALVATNPV SKVVSQFIHL YLMGCNPFWM LCEGIYHLTL 240
 IVVAVFAEK HLMWYFLGW GPPLIPACIH AIARSLYND NCWISSDTHL LYIHHGPICA 300
 ALLVNLFFLL NIVRVLTITKL KVTHQAESNL YMKAVRATLI LVPLLGIEFV LIPWRPEGKI 360
 AEEVYDIIMH ILMHFQGLLV STIFCFNGE VQAILRRNWN QYKIQFQNSP SNSEALRSAS 420
 YTVSTISDGP GYSHDCPSEH LMGSIHIDIE NVLLKPENLY N

Seq ID NO: 338 DNA sequence
 Nucleic Acid Accession #: NM_001795
 Coding sequence: 25-2379

1 11 21 31 41 51
 GCAAGATCTG TTCTCTCTGG GAAGATGCAG AGGCTCATGA TGCTCCTCGC CACATCGGGC 60
 GCCTGCCTGG GCCTGCTGGC AGTGGCAGCA GTGGCAGCAG CAGGTGCTAA CCCTGCCCAA 120
 CGGACACACC ACAGCCTGCT GCCACCCAC CGGCGCCAAA AGAGAGATTG GATTGGAAAC 180
 CAGATGCACA TTGATGAAGA GAAAAACACC TCACTTCCCC ATCATGTAGG CAAGATCAAG 240
 TCAAGCGTGA GTGCAAGAA TGCCAAGTAC CTGCTCAAAG GAGAATATGT GGGCAAGGTC 300
 TTCCGGGTGG ATGCAGAGAG AGGAGACGTG TTCCGCAATT AGAGGCTGGA CGGGAGAAAT 360
 ATCTCAGATG ACCACCTCAC TGCTGTCAAT GTGCAACAGG ACACCTGGTA AAACCTGGAG 420
 ACTCCTTCCA GCTTCACCAT CAAAGTTTCA GACGTGAACG ACACTGGCC TGTGTTCAAG 480
 CATCGGTTGT TCAATGCGTC CGTGCTGAG TCGTGGCTG TGGGACCTC AGTCATCTCT 540
 GTGACAGCAG TGGATGCAGA CGACCCCACT GTGGGAGACC ACGCTCTGT CATGTACCAA 600
 ATCCTGAAGG GGAAGAGATA TTTGCCATC GATAATTCTG GACGTATTAT CACATAAAG 660
 AAAAGCTTGG ACCGAGAGAA GCAGGCCAGG TATGAGATCG TGGTGAAGC GCGAGATGCC 720
 CAGGGCTCTC GGGGGACTG GGGCAGGCC ACCGTGCTGG TCACTCTGCA AGACATCAAT 780
 GACAACTTCC CTTCTTCTAC CCAGACCAAG TACACATTTG TCGTGCCTGA AGACACCCGT 840
 GTGGGACCTT CTGTGGGCTC TCTGTTTGTG GAGGACCCAG ATGAGCCCCA GAACCGGATG 900
 ACCAAGTACA GCATCTTGGG GGGCGACTAC CAGGACGCTT TCACCATGTA GACAAAACCC 960
 GCCACACAAG AGGGCATCAT CAAGCCCATG AAGCCTCTGG ATTATGAATA CATCCAGCAA 1020
 TACAGCTTCA TCGTCGAGGC CACAGACCCC ACCATCGACC TCGATACAT GAGCCCTCCC 1080
 GCGGGAAAAC GAGCCCAAGT CATTATCAAC ATCAGATG TGGACGAGCC CCCCATTTC 1140
 CAGCAGCCTT TCTACCACT CCAGCTGAAG GAAAACAGA AGAAGCCTCT GATTGGCACA 1200
 GTGCTGGCCA TGGACCTGTA TCGGCTAGG CATAGCATG GATACTCCAT CCGCAGGACC 1260
 AGTGACAAGG GCCAGTTCTT CCGAGTCACA AAAAAGGGGG ACATTTACAA TGAGAAAGAA 1320
 CTGACACAGG AAGTCTACCC CTGGTATAAC CTGACTGTGG AGGCCAAAGA ACTGGATTCC 1380
 ACTGGAACCC CCACAGGAAA AGAATCCATT GTGCAAGTCC ACATTTGAAT TTTGGATGAG 1440
 AATGACAATG CCGCGGAGTT TGCCAAGCCC TACCAGCCCA AAGTGTGTA GAACGCTGTC 1500
 CATGGCCAGC TGGTCTGCA GATCTCGCA ATAGACAGG ACATAACACC ACGAAACGTG 1560
 AAGTTCAAAT TCACCTTGAA TACTGAGAAC AACITTACCC TCACGGATAA TCACGATAAC 1620
 ACGGCCAACA TCACAGTCAA GTATGGCAG TTTGACCGGG AGCATACCAA GGTCCACTTC 1680
 CTACCGGTGG TCATCTCAGA CAATGGGATG CCAAGTCGCA CGGCACCAAG CACGCTGACC 1740
 GTGGCCGTGT GCAAGTGCAA CGAGCAGGGC GAGTTCACTT TCTGCGAGGA TATGGCCGCC 1800
 CTAGTGGGGT GAGTCATCCA GGCAGTGGTA GCCATCTTAC TCTGATCCTT CACCATCACA 1860
 GTGATCAACC TGCTCATCTT CCGCGGCGG CGGCTCCGGA AGCAGGCCCC CGGCGACGGC 1920
 AAGAGCGTGC CGGAGATCCA CGAGCAGCTG GTCACTACG ACGAGGAGGG CGGCGGCGAG 1980
 ATGGACACCA CCAGCTACGA TGTGTCTGGT CTCAACTCGG TCGCGCCGGG CGGGGCCAAG 2040
 CCGCGCGGGC CCGCGCTGGA CGCCCGGCGT TCCCTCTATG CGCAGGTGCA GAAGCCACCG 2100
 AGGCACCGGC TGGGGGACA CGGAGGGCCC GGGGAGATGG CAGCCATGAT CGAGGTGAAG 2160
 AAGGACGAGG CGGACACAGA CGGCGACGGC CCCCCTACG ACACGCTGCA CATCTACGGC 2220
 TACGAGGGCT CCGAGTCCAT AGCCGAGTCC CTCAGCTCCC TGGGACCGA CTCATCGAC 2280

TCTGACGTGG ATTACGACTT CCTTAAACGAC TGGGACCCCA GGTTTAAGAT GCTGGCTGAG 2340
 CTGTACGGCT CGGACCCCGG GAGGAGGCTG CTGTATTAGG OGGCCGAGGT CACTCTGGGC 2400
 CTGGGGACCC AAACCCCTCG CAGCCAGGC CAGTCAGACT CCAGGSCACCA CAGCCTCCAA 2460
 AAATGGCAGT GACTCCCGAG CCCAGCACCC CTCTCTCGTG GGTCCACAGAG ACCTCATCAG 2520
 5 CTTGGGGATA GCAAACTCCA GGTTCCTGAA ATATCCAGGA ATATATGTCA GTGATGACTA 2580
 TTCTCAAATG CAGCAAACTC CAGGCTGGTG TTCTGTCTGG GCTCAGACAT CCACATAACC 2640
 CTGTCAACCA CAGACCGCGG TCTAACTCAA AGACTTCTCT TGGCTCCCA AGGCTGCAAA 2700
 GCAAAACAGA CTGTGTTTAA CTGCTGCAGG GTCTTTTCTT AGGGTCCCTG AACGCCCTGG 2760
 10 TAAGGCTGGT GAGGTCTCTG TGCCTATCTG CTTGGAGGCA AAGGCTGGA CAGCTTGACT 2820
 TGTGGGGCAG GATTCTCTGC AGCCCATTCG CAAGGAGAGC TGACCATCAT GCCCTCTCTC 2880
 GGGAGCCCTA GCCCTGCTCC AACTCCATAC TCCACTCCAA GTGCCCCACC ACTCCCCAAC 2940
 CCCTCTCCAG GCCTGTCAAG AGGGAGGAAG GGGCCCATG GCAGCTCCTG ACCTTGGGTC 3000
 CTGAAGTGAC CTCAGTGCCG TGCCATGCCA GTAAGTGTGC TGTACTGAGC ACTGAACCAC 3060
 ATTACGGGAA ATGCTTATTA AACCTTGAAG CAAGTGTGAA TTCATTCTGG AGGGCAGTG 3120
 15 GAGATACAGA GTACACAGAT ACAGGGTGAG GGCACCTCC ACACCCACCC CCTCTGGAGA 3180
 AGGCCTGGAA GAGCTAGAGC CTTGCTTGA GACTCCTCAG CACCCCTCCA GTTTTGCCTG 3240
 AGAAGGGCA GATTTCTCCG GAGATCAGAA GACGTCTCCG CTTCTCTGCC TCACCTGGTC 3300
 GCCAATCCAT GCTCTCTTTC TTTCTCTGT CTACTCCTTA TCCCTTGGTT TAGAGGAACC 3360
 20 CAAGATGTGG CCTTTAGCAA AACTGACAAT GTCCAAACCC ACTCATGACT GCATGACGGA 3420
 GCGAGGACAT GTCTCTTACA CCTCGCTGTT GTCACATCT AGGGAACTGA CCCTCAGGCA 3480
 CACCTTGACG AAGGAGGGCC CTGCCCTGCC CAACCTCTGT GGTCAACCAT GCATCATTC 3540
 ACTGGAAAGT TTCACTGCAA ACACACCTTG GAGAAGTGGC ATCAGTCAAC AGAGAGGGGC 3600
 AGGAAGGGAG ACACCAAGCT CACCTTCTGT CATGGACCGA GGTCCCACT CTGGCAAGC 3660
 25 CCCTCACACT GCAAGGATG GTAGATAACA CTGACTTGTG TGTTTTAAAC AATAACTAGC 3720
 TTCTTATAAT GATTCTTTTA CTAATGATAC TTACAAGTTT CTAGCTCTCA CAGACATATA 3780
 GAATAAGGCT TTTTGTATAA TAAGCAGGTT GTTATTTAGG TTAACAATAT TAATTCAGGT 3840
 TTTTATGTG GAAACCAAT TCCTGTAACC TTCTATTTTC TATAATTGTA GTAATTGCTC 3900
 TACAGATAAT GTCTATATAT TGGCCAAACT GGTGCATGAC AAGTACTGTA TTTTATTATA 3960
 30 CCTAAATAAA GAAATATCTT TAGCCTGGGC AACAAAAAAA

Seq ID NO: 339 Protein sequence
 Protein Accession #: NP_001786

1 11 21 31 41 51
 MQRLLMLLAT SGACLGGLAV AAVAAAGANP AQRDTHSLLP THRRQKRDIW MNQMHIDEER 60
 35 NTSLPHEHVK IKSSVSRKNA KYLLKGEYVG KVRVDAETG DVFAIERLDR ENISEYHLTA 120
 VIIVDKDTGEN LETPSSFTIK VHDVNDNWPV FTHRLFNASV PESSAVGTSV ISVTAVDADD 180
 40 PTVGDSHASVM YQILKGEYF AIDNSGRIIT ITKSLDREKQ ARYEIVVEAR DAQGLRGDSG 240
 TATVLVTLQD INDNFFPFTQ TKYTFVVPED TRVGTSVGSL FVEDPDEPQN RMTKYSILRG 300
 DYQDAFTIET NPAHNEGIIK PMKPLDYEYI QQYSFIVEAT DPTIDLRYS PPAGNRAQVI 360
 INITDVDEFP IFQPPFYHFO LKENQKKPLI GTVLAMPDPA ARHSIGYSIR RTSKDGQFPR 420
 VTKKGDYNE KELDREVYPW YNLIVEAKEL DSTGPTPKE SIVQVHIEVL DENDNAPEFA 480
 45 KPYQPKVCEN AVHGQVLVLI SAIDKIDITPR NVKFKPTLNT ENNPTLIDNH DNTANITVKY 540
 GQFDREHTKV HFLPVVISDN GMPSTRGTST LTVAVCKCNE QGEFTFCEDM AAQVGVSIIQA 600
 VVAILLCILT ITVITLLIFL RRLRLKQARA HGKSVPEIHE QLVTYDEEGG GEMDTTSYDV 660
 SVLNSVRRGK AKPRPALDA RPSLYAQVQK PPRHAPGAHG GPGEMAAMIE VKKDEADHDG 720
 DGPFPYDLHI YGEGSESIA ESSLGLTDS SDSVDVDFL NDWGRPFKML AELVGSOPRE 780
 50 ELLY

Seq ID NO: 340 DNA sequence
 Nucleic Acid Accession #: NM_003088
 Coding sequence: 112-1593

1 11 21 31 41 51
 CGCGAGGGTG CGTGGGGGCC GCGGCAGCCG AACAAAGGAG CAGGGGCGCC GCGCAGGGA 60
 60 CCGGCCACCC ACCTCCCGGG GCGCGCAGC GGCCTCTCGT CTAAGTCCAC CATGACCGCC 120
 AACGGCACAG CGAGGCGGTG GCAGATCCAG TTCGGCTCA TCAACTGCGG CAACAAGTAC 180
 CTGACGCGCG AGGCGTTCGG GTTCAAGGTG AACCGCTCCG CCAGCAGCCT GAAGAAGAAG 240
 CAGATCTGGA CGCTGGAGCA GCGCCCTGAC GAGGCGGGCA GCGGGCCGCT GTGCGTGGC 300
 AGCCACCTGG GCGGCTACCT GCGGCGGAC AAGGACGGCA ACGTGACCTG CGAGCGCGAG 360
 65 GTGCCCGGTC CCGACTGCGG TTTCTCATC GTGGCGCAGC ACGACGCTCG CTGCTTCGCG 420
 CAGTCCGAGG CGCACCGGCG CTACTTCGGC GGCACCGAGG ACCGCTCTGC CTGCTTCGCG 480
 CAGACGCTGT CCGCCCGCGA GAAGTGGAGC GTGCACATCG CCATGCACCC TCAGGTCAAC 540
 ATCTACAGTG TCACCGGTAA GCGCTACGCG CACCTGAGCG GCGGCGCGCG CAGAGAGATC 600
 GCGGTGGACC GCGAGCTGCC CTGGGGCGTC GACTCGCTCA TCACCTCTCG CTTCCAGGAC 660
 70 CAGCGCTACA GCGTGAGAGC GCGCGACCC GCGTCTCTGC GCCACGACGG GCGCCTGGT 720
 GCGCGCCCCG AGCCGCGCCAC TGGCTACAGC CTGGAGTTCC GCTCCGCGAA GGTGGCCTTC 780
 CGGAGCTGCG AGGGCGGTTA CCTGGCGCGG TCGGGGCCCA GCGGACGCT CAAGCGGGGC 840
 AAGGCCACCA AGGTGGGCAA GGACGAGCTC TTTGCTCTGG AGCAGAGCTG CGCCAGGTG 900
 GTGCTGCAGG CGGCCAACGA GAGGAACGTG TCCACGCGCC AGGGTATGGA CCGTCTGCCC 960
 75 AATCAGGAGC AGGAGACCGA CCAGGAGACC TTCCAGCTGG AGATCGACCG CGACACCAA 1020
 AAGTGTGCTT TCGTACCCA CACGGGCAAG TACTGGACGC TGACGGCCAC CGGGGGCGTG 1080
 CAGTCCACCG CCTCCAGCAA GAATGCCAGC TGCTACTTTG ACATCGAGTG GCGTGACCGG 1140
 CGCATACAC TGAGGCGCTC CAATGGCAAG TTTGTGACCT CCAAGAGAA TGGGAGCTG 1200
 GCGGCTCTCG TGAGACAGC AGGGGACTCA GAGCTCTCC TCATGAAGCT CATCAACCGC 1260
 80 CCCATCATCG TGTTCGCGGG GGAGCATGGC TTCTACGGCT GCCCGAAGGT CACGGGCAAC 1320
 CTGAGCGCCA ACCGCTCCAG CTATGACGTC TTCCAGCTGG AGTTCAACGA TGGCGCTTAC 1380
 AACATCAAG ACTCCACAG CAAATACTGG ACGGTGGGCA GTGACTCCGC GGTCAACAGC 1440
 AGCGGGACA CTCCTGTGGA CTTCTCTCTG GAGTCTGCG ACTATAACAA GGTGGCCATC 1500
 AAGGTGGGCG GCGCTACCTT GAAGGGCSAC CACGAGGCG TCCTGAAGGC CTGCGCGGAA 1560
 85 ACCGTGGAGC CGGCTCTGCT CTGGGAGTAC TAGGGCCGCG CCGTCTTTC CCGCCCTGCG 1620
 CCAATGGCG GCTCTGCGCA ACCCTCCCTG CTAACCCCTT CTCCGCGAGG TGGGCTCCAG 1680
 GCGGAGGAGC AAGCCCTCTT GCCTTTCAAA CTGGAAACCC CAGAGAAAC GGTGCCCCCA 1740
 CCTGTCCGCC CTATGAGTAC CCCACTCTCC CCTCGCGCG GGTTCCTTAC TCCGCTCGGG 1800

TCAGCGGCTG CGGCGTGGCC CTGGGAGGGA TTTCAGATGC CCTGCCCCC TTGCTGCCA 1860
 CGGGGCGAGT CTGGCACCTC TTCTTCTGA CCTCAGACGG CTCTGAGCT TATTCTCTG 1920
 GAAGCGGCTA AGGAGCGGTG GGGGCTGGG AGCCCTGGGC GTGTAGTGA ACTGGAATCT 1980
 TTTGCTCTC CCAGCCACT CTCTCCAGCC CCCCAGGAGA GCTGGGCACA TGTCCTCAAGC 2040
 CTGTCACTGG CCTCCTCTG TGCACGTGCC CGAAACCCG TGCTTGGGA GGAAGCTGT 2100
 CGGAGGGCT AGGACTGACC CTGTGTGTGT TTTTGTGGT GGTGGCTGA AACAGCCCT 2160
 CTCCCACTG GGAGAGGCTC AGCCTGGCTC CTCTCCCTG AGCGGAGGG CGTGACGGCC 2220
 ACAGGGTCT CCGCTGCAC GTTCTGCCAA GGTGGTGGT GCGGGCGGGT AGGGGTGTGG 2280
 GGGCGCTCT CCTCCTCTC TTTCTCTTC ACCCTAGCT GACTGGAAGC AGAAATGAC 2340
 CAAATCAGTA TTTTCTTAA TGAAATATTA TTGCTGGAG CGTCCAGGC AAGCCTGGCT 2400
 GTAGTAGCA GTGATCTGGC GGGGGGCTC TCAGCACCT CCCCAGGGG TGCTCTCAG 2460
 CCCCCTCTT CGTCTCTCC CGTCCAGCC CAGCCTGGG CTGGGCTGC CGACACCTG 2520
 GCCAGAGCC CTGCTGTGAT TGGTGTCTCC TGGGCTCTCC GGGTGGATGA AGCCAGGCT 2580
 CGCCCTCTC GGGAGCCCTC GGGTGAAGC CCGGGGCCCC CTGCTGCCA GCCTCCCCG 2640
 TCCCAACAT GCATCTCACT CTGGGTGTCT TGGTCTTTA TTTTGTGTA GTGTCAATTG 2700
 TATAACTCTA AACGCCATG ATAGTAGCTT CAACTGGAA ATAGCGAAAT AAAATAACTC 2760
 AGTCTGC

Seq ID NO: 341 Protein sequence
 Protein Accession #: NP_003079

1 11 21 31 41 51
 MTANGTAEAV QIQGLINCG NKYLTAFAFG FKVNASASSL KKKQIWTLEQ PPDEAGSAAV 60
 CLRSHLGRYL AADKDGIVTC EREVPGPDCR FLIVAHDDGR WSLQSEARRR YPGGTEDRLS 120
 CPAQTVSPAE KMSVHIAMHP QVNIYSVTRK RYAHLSARPA DEIAVDREVP WGVDSLITLA 180
 FQDQRYSVQT ADHRFLRHDG RLVARPEPAT GYTLERFSKG VAFRDCEGRY LAPSGPSGTL 240
 KAGKATKVGK DELPALQSC AQVVLQAANE RNVSTRQMD LSNQDEETD QSTFQLEIDR 300
 DTKKCAPRTH TGKWTLTAT GGVQSTASSK NASCYFDIEW RDRRITLRAS NGKFVTSKKN 360
 GQLAASFCAT GDSLEFLMKL INRPIIVFRG EHGFIGCRKV TGTLDANRSS YDVFQLEFND 420
 GAYNIKSDTG KYWTVGSDSA VTSSGDTFVD PFPEFCYNNK VAIKVGGRYL KGDHAGVLKA 480
 SAETVDPASL WEY

Seq ID NO: 342 DNA sequence
 Nucleic Acid Accession #: FGENESH predicted
 Coding sequence: 660..1705

1 11 21 31 41 51
 CGCTCCGAC ACATTTCTGT TCGCGGCTA AGGAAACTG TTGGCCGCTG GGCCCGCGGG 60
 GGGATTCTT GCAGTTGGGG GGTCCGTCGG GAGCGAGGGC GGAGGGGAAG GGAGGGGGA 120
 CCGGTTGGG GAAGCCAGCT GTAGAGGGCG GTGACCGCGC TCAGACACA GCTCTGCTC 180
 CTCGAGCGGG ACAGATCCAA GTTGGGAGCA GCTCTGCGTG CCGGGCTCA GAGAATGAGG 240
 CGCGGTTGG CCTCTGGCAG CGCTCTGGC CCGGGCCGGG CGCGCGGAA 300
 CACCCACTG CCGACCTGTC TGGCTGCTCG GCCTCGGGGG CTGTCTACAG CCTGCACCAC 360
 GCTACCATGA AGCGGCAGGC GCGCGAGGAG GCCTGCATCC TGCGAGGTGG GCGCTCAGC 420
 ACCGTGCTG CCGGCGCGCA GCTGCGCGCT GTGCTCGCGC TCCTGCGGGC AGGCCACGG 480
 CCGGAGGGG GCTCCAAAGA CCTGCTGTTT TGGGTGCGAC TGGAGCGCAG GCGTTCACCA 540
 TGCACTCTG AGAACGAGCC TTTGCGGGGT TTCTCCTGGC TGCTCTCGA CCGCGCGGT 600
 CTCGAAAGCG ACACCTGTC GTGGGTGGAG GAGCCCCAAC GCTCTGCAC CGCGCGGAGA 660
 TGCGCGGTAC TCCAGGCCAC CGGTGGGGTC GAGCCCGCAG CTGGAAGGAG ATGCGATGCC 720
 ACCTGCGGCG CAACGGCTAC CTGTGCAAGT ACCAGTTTGA GGTCTTGTGT CTGCGCGGC 780
 GCGCGCGGG CCGCTCTAAC TTGAGCTATC GCGCGCCCTT CCAGCTGCAC AGCGCGGCTC 840
 TGGACTTCA TCCACTTGG ACCGAGGTGA GTGCGCTCTG CCGGGACAGC CTCCGATCT 900
 CAGTTACTTG CATCGCGGAG GAAATCGGCG CTCGCTGGGA CAAACTCTG GCGCATGTGT 960
 TGTGTCCCTG CCGCGGAGG TACCTCCGTG CTGGCAATG CGCAGAGCTC CTAAGTCTCC 1020
 TAGACGACTT GGGAGGCTTT GCCTGCGAAT GTGCTACGGC CTTGAGCTG GGAAGGACG 1080
 GCGGCTCTG TGTGACCACT GGGGAAGGAC AGCCGACCTT TGGGGGAGC GGGGTGCCA 1140
 CAGGCGGCT GCGCGGCACT GCAACGAGCC CCGTGCGCGA GAGAACATGG CCAATCAGGG 1200
 TCGACGAGAA GCTGGGAGAG ACACCACTT TCCCTGAACA AGACAATTCA GTAACATCTA 1260
 TTCTTGAGAT TCCTCGATGG GGATCAGAG GCAAGATGTC TACCCTTCAA ATGTCCCTTC 1320
 AAGCGAGTC AAAGGCCACT ATCACCCTAT CAGGAGCGGT GATTTCAGG TTATATCTA 1380
 CGACTCTCT TGCCACTCTT CAGGCTTTCG ACTCTCTCT TGCCGTGCTC TTCAATTTG 1440
 TGAGCACAGC AGTAGTAGTG TTGGTATCTT TGACCATGAC AGTACTGGGG CTGTCAAGC 1500
 TCTGCTTCA CGAAAGCCCC TCTTCCAGC CAAGGAAGGA GTCTATGGGC CCGCGGGCC 1560
 TGGAGAGTGA TCCTGAGCCC GCTGCTTTGG GCTCCAGTTC TGCAATTG ACACAATG 1620
 GGGTGAAAGT CCGGAGCTGT GATCTGCGGG ACAGAGCAGA GGTGCTCTG CTGGCGGAGT 1680
 CCCCCTCTG CTCTAGTATG GCATAG

Seq ID NO: 343 Protein sequence
 Protein Accession #: FGENESH predicted

1 11 21 31 41 51
 MGKDFMTKP KAFATKAKID KNDLIKLSF CTAKETIIRV NSOPTDWQKT FAIYPSDKGV 60
 IARIYKELEQ IYKKKPKTK LRTHFLSRPK GNCNPLGPRG DSWQLGGPSG ARAEGKGGT 120
 GLGKPAVEGG DRAPDTALRP RAGQIQVGS SAGGASENEA GVRPVPPLAG ALARAGRRT 180
 PHCRPCMLLG LGGLLQAPAP YHEAAGGRGG LHPARWGAQH RACGRRAARC ARAPAGRPR 240
 RRGLQRPAVL GRTGAQAPFL HPGERAFAGF LLAVLRPRRS RKRHAAGVGG APTLLHRAEM 300
 RGTGHRWRG ARSWKEMRCH LRANGYLCKY QFEVLCPAPR PGAASNLVSYR APFQLHSAAL 360
 DFSPPGTEVS ALCRQLPLIS VTCIADEIGA RWDKLSGDVL CPCPGRYLRA GKCAELPNCL 420
 DDLGGFACEC ATGFELGKDG RSCVTSSEGG PTLGGTGVPRT RRPPATATSP VPQRTWPIRV 480
 DEKLGETPLV PEQDINSVSI PEIPRWGSQS TMTSLQMSLQ AESKATITPS GSVISKFNST 540
 TSSATPQAFD SSSAVVFIFV STAVVVLVIL TMTVLGLVLK CFHESPSSQP RKESMGPPGL 600
 ESDPEPAALG SSSAHCNNG VKVGDCDLRD RAEGALLAES PLGSSDA

Seq ID NO: 344 DNA sequence
Nucleic Acid Accession #: NM_012072
Coding sequence: 149-2107

5
0
5
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
AAAGCCCTCA GCCTTTGTGT CCTTCTCTGC GCGGAGTGG CTGAGCTCA CCCCAGCT 60
CCCCTTGGGG CCCAGCTGGG AGCCGAGATA GAAGCTCCTG TCGCCGCTGG GCTTCTCGCC 120
TCCCGCAGAG GGCACACAG AGACCGGGAT GGCACCTCC ATGGGCTGTC TGCTGCTGCT 180
GCTGCTGCTC CTGACCCAGC CCGGGGCGGG GACGGGAGCT GACACGGAGG CCGTGGTCTG 240
CGTGGGACC GCCTGTCTACA CGGCCACTC GGGCAAGCTG AGCGCTCGCG AGGCCAGAA 300
CCACTGCAAC CAGAACGGGG GCAACCTGSC CACTGTGAAG AGCAAGGAGG AGGCCAGCA 360
GTCACAGCGA GTACTTGCCC AGCTCCTGAG GCGGAGGCA GCCCTGACGG CGAGGATGAG 420
CAAGTCTCTG ATTGGGCTCC AGCGAGAGAA GGGCAAGTGC CTGGACCTTA GTCTGCCGCT 480
GAAGGGCTTC AGCTGGGTGG GCGGGGGGGA GGACACGCT TACTCTAACT GGCACAAAGGA 540
GCTCCGGAAC TCGTGCATCT CCAAGCGCTG TGTGTCTCTG CTGCTGGACC TGTCCAGGCC 600
GCTCCTTCCC AACCGCCTGC CCAAGTGGTC TGTGTGCAAA GTTCAGCTTC AAAGCATGT GCCGCTCTC 660
CGGAAAGTAA ATTGAGGGCT TGTGTGCAAA GTTCAGCTTC AAAGCATGT GCCGCTCTC 720
GGCCCTGGGG GGGCCAGGTC AGGTGACCTA CACCACCCCC TTCCAGACCA CCAGTTTCCTC 780
CTTGGAGGCTT GTGCGCTTTG CCTCTGCGGC CAATGTAGCC TGTGGGGAAG GTGACAAAGGA 840
CGAGACTCAG AGTCATTATT TCCTGTGCAA GGAGAAGGCC CCCGATGTGT TCGACTGGGG 900
CAGCTCGGGC CCCCTCTGTG TCAGCCCCAA GTATGGCTGC AACTTCAACA ATGGGGGCTG 960
CCACCAGGAC TGCTTTGAAG GGGGGGATGG CTCCTTCCCT TCGGCTGCC GACCAGGATT 1020
CGGCTGCTG GATGACCTGG TGACCTGTGC CTCTGAAAC CCTTGCAGCT CCAGCCCATG 1080
TCGTGGGGG GCGACGTGCG TCCTGGGACC CCATGGGAAA AACTACAGT GCGCTGCCCT 1140
CCAAGGGTAC CAGCTGGAAT CGAGTCAGCT GGAATGTGTG GAGCTGGATG AATGCCAGGA 1200
CTCCCCCTGT GCCCAGGAGT GTGTCAACAC CCCTGGGGGC TTCCGCTGG AATGCTGGGT 1260
TGGCTATGAG CCGGGGCTGT CTGGAGAGGG GGCTGTCTAG GATGTGGATG AGTGTGCTCT 1320
GGGTGCTGCG CCTTGGCCCC AGGGCTGCAC CAACACAGAT GGCTCATTTT ACTGCTCTCT 1380
TGAGAGGGCG TACCTCTGG CCGGGGAGGA CGGACTCAG TGCCAGGACG TGGATGAGTG 1440
TGTGGGCCCC GGGGGCCCCC TCTGGCAGAG CTGTGTCTTC AACACACAAG GGTCTCTCCA 1500
CTGTGGCTGC CTGCCAGGCT GGGTGTCTGC CCAAAATGGG GTCTCTTGA CCAATGGGGC 1560
TGTGTCTCTG GGACCAACAT CTGGGCCCCC CGATGAGGAG GACAAGGAG AGAAGGAAGG 1620
GGCTACACCC ACCACAAGTA GACCTTCGCT GTCATCTGAC GCCCCATCA CATCTGCCCT 1740
ACTCAAGATG CTGGCCCCCA GTGGGTCTTC AGGCGTCTGG AGGGAGCCCA GCATCCATCA 1800
CGCCACAGCT GCCTCTGGCC CCCAGGAGCC TGCAAGTGGG GACTCTCCG TGGCCACACA 1860
AAACAAAGAT GGCACCTGAG GGCACAAAGCT GCTTTTATTC TACATCTAG GCACCGTGGT 1920
GGCCATCTTA CTCTGCTGG CCTTGGCTCT GGGGCTACTG GTCTATCGCA AGCGGAGAG 1980
GAAGAGGGAG GAGAAGAAGG AGAAGAAGCC CCAGAAATGG CGAGACAGTT ACTCCTGGGT 2040
TCCAGAGCGA GCTGAGAGCA GGGCCATGGA GAACCAATAC AGTCCGACAC CTGGGACAGA 2100
CTGCTGAAGG TGAGGTGGCC CTAGAGACAC TAGAGTCACC AGCCACCATC CTCAGAGCTT 2160
TGAACCTCCC ATTCCAAGG GGCACCCACA TTTTGTGAA AGACTGGACT GGAATCTTAG 2220
CAACCAATTG TAAGTCTCCT CCTTAAAGGC CCTTGGAAAC ATGCAGGTAT TTTCTAGG 2280
TGTTTGATGT TCTGGAAGTG GAAGCTGTGT GTTGGCTGCG CACGGTGGGG ATTTCTGTAC 2340
TCTATAATGA TTGTACTCC CCTTCCCTTT TCAAATCCA ATGTGACCAA TTCCGGATCA 2400
GGGTGTGAGG AGGCTGGGGC TAAGGGGCTC CCTGAATAT CTCTCTGCT CACTTCCACC 2460
ATCTAAGAGG AAGAGGTGAG TTGCTCATGC TGATTAGGAT TGAATGATT TGTTCCTCTT 2520
CCTAGATGTA AACTTAATAT AATTAATAT TCAATTAGT AAGAAGATCT GGTTTTTTGT 2580
TCAAAGGGAA CATGTTGGGA CTGGAACAT TTCTTTACAT TTGCATTCTT CCATTTCCGC 2640
AGCACAGTCT TTGCTAAATG TGATCTGTT GACATCTCTC AGAATGGCCA GAAGTGCAAT 2700
TAACCTCTTA GGTGGCAAGG AGGCAGGAAG TGCTCTTTA GTTCTTACAT TTCTAATAGC 2760
CTTGGGTTTA TTTGCAAAAG AAGCTTGAAG AATATGAGAA AAGTTGCTG AAGTGCAATTA 2820
CAGGTGTTTG TGAAGTCACA TAATCTACGG GGCTAGGGCG AGAGAGGCCA GGGATTGT 2880
CACAGATACT TGAATTAATT CATCCAAATG TACTGAGGT ACCACACAT TGACTACGGA 2940
TGTGATCAAC ACTAACAGG AAACAAATTC AAGGACAAAC TGCTTTGAG CCAGGGCAGG 3000
CCTCAGACAC CTTGCTCTGC CCGCCGCTC CACTTCAATC TGCCCGGAAT GCCAGTGCTC 3060
CGAGCTCAGA CAGAGGAAGC CCTGCAGAAA GTTCCATCAG GCTGTTTCTT AAAGGATGTG 3120
TGAAAGGGAG ATGATGACT GTGTTTGAAG AGTGTCTATT TTAAGCATT TTAGCACAGT 3180
TCATAGTCCA CAGTTGATG AGCATCTGGA GATTTTAAAT CCTGAAGTGT GGGTGGCGCA 3240
CACACCAAGT AGGGAGCTAG TCAGGCAGTT TGCTTAAGGA ACTTTGTTC TCTGTCTCTT 3300
TTCTTAAATA TTTGGGGTAA GGAGGGGAA AAGAGGGAAG AGATGACTA ACTAAATCA 3360
TTCTTAAATA TTTGGGGTAA GGAGGGGAA AAGAGGGAAG AGATGACTA ACTAAATCA 3420
TTTTTACAGC AAAAAGTCTT CAAGGCCATT TAAATATAT CCTCATTTTA AAGTTACAT 3480
TTGCAATAT TTCTCCCTAT GATAATGAG TCGATAGTGT GCATCTTTT TCTCTCTCTC 3540
TCTCTCTCAC ACACACACAC ACACACACAC ACACACACAC CAGAGTCACT AGAAGTAC 3600
CCTGGGGCAC TGGAAACAT TCCTGGGGGT CATGTCTCTT GTGGGCTTTT TACCACCACT GTGCAGGAGA 3660
TGAGTATCTC TGGGAGGCC CTCTCCCAAG GCCCCAAAGC CTCAGAGAAA GGGTGTCTCT 3720
ACAGACAGAG GAAATGTGTC TCCCTCCAAAG AGGTGACACT CTGGAGTGGT TGAAGGGCCA 3780
GGTTTTGCCT TAGCAATGCA TCGGTCTCTG AGGTGACACT CTGGAGTGGT TGAAGGGCCA 3840
CAAGGTGCAG GGTTAATACT CTTGCCAGTT TTGAAATATA AGATGGAGGT TTTGTGCGGC 3900
TTTTAATAGA AACTTAAAGG GGCAGGGGAA GTGAAAGGAA AGTCTCCAGT TTTCAAGTTG 3960
TCGATGGGGC ATTTGGAAC TCTTTTAAAG AAGGGAGACA AAGGCTGTGT CCATTTGGCA AAATCTCCTT 4020
AACTCTGGTG TTTAACAATT AAGGGAGACA AAGGCTGTGT CCATTTGGCA AAATCTCCTT 4080
GGCCACGAGA CTCTAGGTGA TGTGTGAAGC TGGGCGAGTCT GTGGTGTGGA GAGCAGGCAT 4140
CTGTCTGGCC ATTCAGAGGA TTCTAAAGAC ATGGCTGGAT GCGCTGCTGA CCAACATCAG 4200
CACTTAAATA AATGCAAAAT CAACATTCTT CCTCTGGGC CTTGAAATAT CTTGCCCTTA 4260
TCATTTGGGG TGAAGGAGAC ATTTCTGTCT TTGGCTTCCC ACAGCCCAAA CGCAGTCTGT 4320
GTATGATTCC TGGGATCCAA CGAGCCCTCC TATTTTCAAC GTGTTCTGAT TGCTCTCACA 4380
GCCCAGGCC ATGCTGTGTT CTCTGAATGC AGCCCTGTTC TCAACACAGC GGAGGTCTAG 4440
GAACCCCTCT GTGGAAACCA CAAGGGGAGA AATGGGTGAT AAAGAATCCA GTTCTCAA 4500
ACCTTCCCTG GCAGGCTGGG TCCCTCTCCT GCTGGGTGGT GCTTCTCTT GCACACCACT 4560
CCCACACGG GGGGAGAGCC AGCAACCCAA CCAGACAGCT CAGGTGTGTC ATCTGATGGA 4620
AACCACTGGG CTCAAACAG TGCTTTATTC TCTGTTTAT TTTGCTGTT ACTTTGAAGC 4680
ATGGAATTC TTTTGGGG GATCTTGGG CTACAGTAGT GGGTAAACCA ATGCCACCG 4740
GCCAAGAGGC CATTAAACAA TCGTCTTGT CCGTGGGGC CCCAGCTTGC TCGGCGTGG 4800
CCAGTGGGG AATCCAAGG TCACAGTATG GGGAGAGGTG CACCTGCCA CCTGCTAAT

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75

TCTCGCTAGA CACAGTGTCT CTGCCAGGT GACCTGTTC GCAGCAGAAC AAGCCAGGGC 4860
CATGGGGAAG GGGGAAGTTT TCACCTGGAG ATGGACACCA AGACAATGAA GATTGTGTGT 4920
CCAAATAGGT CAATAATTCT GGGAGACTCT TGGAAAAAC TGAATATATT CAGGACCAAC 4980
TCTCTCCCTC CCCTCATCCC ACATCTCAAA GCAGACAATG TAAAGAGAGA ACATCTCACA 5040
CACCCAGCTC GCCATGCCTA CTCATTCCCTG AATTTCAGGT GCCATCACTG CTCCTTCTTT 5100
CTTCTTTGTC ATTTGAGAAA GGATGCAGGA GGACAATTCC CACAGATAAT CTGAGGAATG 5160
CAGAAAAACC AGGGCAGGAC AGTTATCGAC AATGCATTAG AACTTGGTGA GCATCCTCTG 5220
TAGAGGAGCT CCACCCCTGC TCAACAGCTT GGCTTCCAGG CAGACCAAC CACATCTGCT 5280
CTCTGCCTTC GGTGGCCAC ACACCTAAGC GTCATCGTCA TTGCCATAGC ATCATGATGC 5340
AACACATCTA CGTGTAGCAC TACGACGTTA TGTTTGGGTA ATGTGGGGAT GAACTGCATG 5400
AGGCTCTGAT TAAGGATGTG GGGAAAGTGG CTGCGGTGAC TGTGGGCTT GCAAGGCCAC 5460
CTGGAGGCTT GTCTGTAGC CAGTGGTGA AGAGCAAGGC TTCAGGAAGG GCCAGCCACA 5520
TGCCATCTTC CTCTGATCA GGCAAAAAAG TGAATTAAG AAGTCAAAAC TTTATATGCA 5580
TGTGTTATGT CCAATTTGCA GGATGAAGT AGTTTAAAAA AATTTTCTT TCTCTTCAAG 5640
TGTCTTGTCT TTTTCCATCC TCATCACAAG CCCTTGTGTG AGTGTCTTAT CCCTGAGCAA 5700
TCTTTCGATG GATGGAGATG ATCATTAGGT ACTTTTGTCT CAACCTTTAT TCCTGTAAAT 5760
ATTCTCTGTA AAGCTAGTGA AACAGAGATG AGATTGACA AAAAAAATG GAATTAATAA 5820
TAACACAGTC TTTTAAATAC TAACATAGGA AAGCCTTTCC TATTATTCTT CTCTTAGCT 5880
TCTCCATTGT CTAATCAGG AAAACAGGAA AACACAGCTT TCTAGCAGCT GCAAAATGGT 5940
TAAATGCCCC CTACATATTT CCATCACCTT GAACAATAGC TTAGCTTGG GAATCTGAGA 6000
TATGATCCCA GAAACATCTT GTCTCTACTT CGGCTGCAAA ACCATGTTT TAAATCTATA 6060
TGGTTTGTGC ATTTTCTGCA CTAATAATAG AGATGATAAT CCGAATCTCT CATATATCA 6120
CTAATCAAG ACACATATTT CATACTAGAT TCCTGAGACA AATACTCACT GAAGGGCTTG 6180
TTTAAAAATA AATTGTGTTT TGGTCTGTTT TTGTAGATAA TGCCCTTCTA TTTTAGGTAG 6240
AAGCTCTGGA ATCCCTTAT TGTGCTGTGT CTCTTATCTG CAAGGTGGCA AGCAGTTCTT 6300
TTCAGCAGAT TTTGCCACTT ATTCTCTGTA GCTGAAGTTC TTGTCATAGA TTTGGCTTAA 6360
GCTTGAATTA GATCCCTGCA AAGGCTTGCT CTGTGATGTC AGATGTAATT GTAAATGTCA 6420
GTAATCACTT CATGAATGCT AATGAGAAT GTAAGTATTT TTAATGTGT GTATTTCAA 6480
TTTGTGTGAC TAATTCTGGA ATTACAAGAT TTCTATGCAG GATTACCTT CATCCTGTGC 6540
ATGTTTCCCA AACTGTGAGG AGGGAAGGCT CAGAGATCGA GCTTCTCTC TGAGTTCTAA 6600
CAAAATGGTG CTTTGAAGGT CAGCCTTAG GAAGGTGCAG CTTTGTGTC CTTTGAAGCT 6660
TCTGTTATGT GCCTATCCTA ATAACTCTT AAACACATT

Seq ID NO: 345 Protein sequence
Protein Accession #: NP_036204

1 11 21 31 41 51
MATSMGLLLL LLLLTLPQGA GTGADTEAVV CVGTACTYAH SGKLSAAEAQ NHCNQNGNL 60
ATVKSKEEAQ HVQRLVLAQL RREAALTARM SKFWIGLQRE KGKCLDPSLP LKGSFVWGGG 120
EDTPYNNMHK ELRNSCISKR CVSLLLDLSS PLLNRLPKW SEGPGSPGS PGSNIEGFVC 180
KFSFKMCRP LALGGPGQVT YTFPQTSS SLEAVFPASA ANVAGGEGDK DETQSHYFLC 240
KEKAPVDFW GSSGPLCVSP KYGCNFNNGG CHQDCFEGSD GSFLGCRPG FRLLDDLVTCT 300
ASRNPSSSP CRGGATCVLG PHGKNYTCRC PQGYQLDSSQ LDCVDVDEQD DSPCAQSCVN 360
TPGFFRCBW VGYEPGGPGE GACQDVDECA LGRSPCAQGC TNDGSPHCS CEEGYVLAGE 420
DGTQCDVDE CVGPGGLDCL SLCFNTQGSF HCGCLPGWVL APNGVSCIMG PVSLGPSPSG 480
PDEEDKGEKE GSTVPRATA SPTRGPEGTP KATPTTSRPS LSSDAPITSA PLKMLAPSGS 540
SGVWREPSIH HATAASGPQE PAGGDSSVAT QNNDGTDGQK LLLFYILGTV VAILLLALA 600
LGLLVYKRR AKREEKKEKK PQNAADSYSW VPERAESRAM ENQYSPPTGT DC

Seq ID NO: 346 DNA sequence
Nucleic Acid Accession #: Z31560
Coding sequence: <1-966

1 11 21 31 41 51
CACAGCGCCC GCATGTACAA CATGATGGAG ACGGAGCTGA AGCCGCCGGG CCGGCAGCAA 60
ACTTCGGGGG GCGCGGGCGG CAACTCCACC GCGCGGGCGG CCGCGGGCAA CCAGAAAAAC 120
AGCCCGGACC GCGTCAAGCG GCCCATGAAT GCCTTCATGG TGTGTCCTCG CGGCGAGCGG 180
CGCAAGATGG CCCAGAGAGG CCCCAAGATG CACAATCTCG AGATCAGCAA GCGCCTGGGC 240
GCCGAGTGGA AACTTTTGTG GGAGACGGAG AAGCGGCGGT TCATCGACGA GGCTAAGCGG 300
CTGCGAGCGC TGACATGAA GGAGCACCCG GATTATAAAT ACCGGCCCCG GCGGAAAAAC 360
AAGACGCTCA TGAAGAAGGA TAAGTACAAG CTGCCCGGGG GGCTGCTGGC CCGCGGGCGG 420
AATAGCATGG CGAGCGGGGT GCGGGTGGGC GCGGCGCTGG GCGCGGGCGT GAACCAAGCG 480
ATGSACAGTT ACAGCGCATG GAAAGCGCTG AGCAACGGCA GCTACAGCAT GATGCAGGAC 540
CAGCTGGGCT ACCCGCAGCA CCGCGGCTCT AATGCGCACG GCGCAGGCGA GATGCAGCCC 600
ATGCACCGCT ACAGCTGAG GCCTCTGAG TACAATCCA TGACCACTG CAGACCTAC 660
ATGAACGGCT CGCCACCTTA CAGCATGTCC TACTCGCAGC AGGGCACCCC TGGCATGGCT 720
CTTGGCTCCA TGGGTTCTGT GGTCAAGTCC GAGGCCAGCT CCAGCCCCCT TGTGTTTACC 780
TCTTCTCTCC ACTCCAGGGC GCGCTGCCAG GCGGGGACCT TCGGGACAT GATCAGCATG 840
TATCTCCCCG GCGCGGAGGT GCGGGAACCC GCGGCCCCCA GCAGACTTCA CATGTCCAG 900
CACTACCCAGA GCGGCGCGGT GCGCGGACCG GCCATTAAAG GCACACTGCC CCTCTCACAC 960
ATGTAGAGGC CGGACAGCGA ACTGGAGGGG GGAGAAATTT TCAAGAAAAA ACGAGGGAAA 1020
TGGAGGGGTT GCAAAAGAGG AGAGTAAGAA ACAGCATGGA GAAACCCCGG TACGCTCAAA 1080
AAAAA

Seq ID NO: 347 Protein sequence
Protein Accession #: CAA83435

1 11 21 31 41 51
HSARMYNNME TELKPPGPQO TSGGGGNGST AAAAGGNQKN SPDRVKRPMN AFMVWSRGQR 60
RKMAQENPM HNSEISKRLG AEWKLLSETE KRPPIDEAKR LRALHMKRHP DYKYRPRRT 120
KTLMKDKYT LPGLLAPGG NSMASGVGVG AGLGAGVNQR MDSYAHNMGW SNGYSMMQD 180
QLGYPOHPLG NAHGAAQMP MHRVDVSLQ YNSMTSSQTY MNGSPYYSMS YSQGTGPGMA 240
LGSMSGVVKS EASSSPFVVT SSSHSRAPCQ AGDLRDMISM YLPGAEPFEP AAPSRLEMSQ 300
HYQSGVPVPT AINGTLPLSH M

Seq ID NO: 348 DNA sequence
Nucleotide Accession #: NM_002638
Coding sequence: 120-473

```
1      11      21      31      41      51
|      |      |      |      |      |
CAATACAGCT AAGGAATTAT CCTTGTAAA TACCACAGAC CGCCCTTGA GCCAGGCCAA 60
GCTGGACTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120
TGAGGGCCAG CAGCTTCTTG ATGTGGTGG TGTCTCTCAT CGCTGGGAAG CTGGTTCTAG 180
AGGCAGCTGT CACGGGAGTT CCTGTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCAT 240
TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAGTTAA AGGTCAAGAT AAAGTCAAAG 300
CGCAAGAGCC AGTCAAAGGT CCAGTCTCCA CTAAGCTTGG CTCCTGCCCC ATTATCTTGA 360
TCCGTTGCGC CATGTTGAAT CCCCCTAACC GCTGCTGAA AGATACTGAC TGCCCAAGAA 420
TCAAGAAGTG CTGTGAAGGC TCTTGGCGGA TGGCTGTGTT CGTTCCCCAG TGAAGGGAGC 480
CGGTCTTTCG TGCACCTGTG CGTCCCAG AGCTACAGGC CCATCTGCTT CTAAGTCCC 540
TGCTGCCCTT CCGCTTCCCA CACTGTCCAT TCTTCTCCCT ATTCAGGATG CCAACGGCTG 600
GAGCTGCTTC TCTCATCCAC TTTCATAATA A
```

Seq ID NO: 349 Protein sequence:
Protein Accession #: NP_002629

```
1      11      21      31      41      51
|      |      |      |      |      |
MRASSFLIVV VFLIAGTLVL EAAVTGVFVK QDVTVKRVP FNGQDPVKQ VSVKQGDVKV 60
AQEPVKGPVS TRPGSCPILL IRCAMLNPPN RCLKDITDCPG IKKCEGSCG MACFPVQ
```

Seq ID NO: 350 DNA sequence
Nucleic Acid Accession #: NM_007183
Coding sequence: 75-2468

```
1      11      21      31      41      51
|      |      |      |      |      |
GAATTCGGGA CAGGACGTGA AGATAGTTGG GTTTGGAGGC GGCCGCCAGG CCCAGGCCCG 60
GTGGAACCTGC CGCCATGCAG GACGGTAACT TCCTGCTGTC GGCCCTGCAG CCTGAGGCCG 120
GCCTGTGCTCT CCGGCGCCTG CCTCTGACC TGCACTGGA CGCCCGGGGC GCGAGGGGGC 180
CGGAGGGCGA GCGGCTCGCG GCAGCCCGCG TCCAGGAGCA GGTCCGCGCC CGCTCTTTGC 240
AGCTGGGACA GCAGCCCGCG CACAACGGGG CCGCTGAGCC CGAGCCTGAG GCCGAGACTG 300
CCAGAGGCAC ATCCAGGGGG CAGTACCACA CCTGCGAGG TGGCTTCAGC TCTGCTCTC 360
AGGGCCTGAG TGGGACAAG ACCTCGGGCT TCGCCGCCAT CGCCAAAGCG GCCTACAGCC 420
CAGCTCCTGT GTCTCTCCGC TCGCCGTGG ATCTGAGCTG CAGTCGGAGG CTGAGTTTCA 480
CCCAATATGG GGGCAGCGCC TTTGGGGCGG CTGGGTACGG GGGTGCCAG CCCACCCCTC 540
CCATGCCAC CAGGCCCGTG TCCTTCCATG AGCGCGGTGG GGTGGGAGC CGGGCCGACT 600
ATGACACACT CTCCCTGCGC TCGCTGCGGC TGGGGCCCGG GGGCCTGGAC GACCGCTACA 660
GCCTGTGTTC TGAGCAGCTG GAGCCCGCG CCACCTCCAC CTACAGGGCC TTTGCGTACG 720
AGCCGCCAGC CAGCTCCAGC TCCAGCCGG CAGGGGGGCT GSACTGGCCC GAGGCCACTG 780
AGGTTTCCCC GAGCCGGACC ATCCGTGCCC CTGGCTGCGG GACCTGCAG CGATTCCAGA 840
GCAGCCACCG GAGCCCGGG GTAGGCGGGG CAGTGCCCGG GGCCGTCTTG GAGCCAGTGG 900
CTCGAGCGCC ATCTGTGCGC AGCCTCAGCC TCAGCCTGGC TGACTCGGGC CACCTGCGGG 960
ACGTGCATGG GTTCAACAGC TACGGTAGCC ACCGAAACCT GCAGAGACTC AGCAGCGGTT 1020
TTGATGACAT TGCCTGCGC TCAGCAGTCA AGTACCTCAT GCTTCAGAC CCCAACCTGC 1080
AGGTGCTGGG AGCGCCCTAC ATCCAGCACA AGTGCTACAG CGATGCAGCC GCCAAGAAGC 1140
AGGCGCGCAG CCTTCAGGCC GTGCTAGGC TGGTGAAGCT CTTCAACCAC GCCAACCAGG 1200
AAGTGACAGC CAGTGCACCA GGTGCCATGC GCAACCTCAT CTAAGACAAC GCTGACAACA 1260
AGCTGGCCCT GGTGGAGGAG AACGGGATCT TCGAGCTGCT GCGGACACTG CGGGAGCAGG 1320
ATGATGAGCT TCGCAAAAT GTACAGGGA TCCTGTGGA CCTTTCATCC AGCGAACACC 1380
TGAAGGACCG CCTGSCCAGA GACACGCTGG AGCAGCTCAC GSACTGGTG TTGAGCCCCC 1440
TGTGCGGGCG TGGGGGTCCC CCCCTCATCC AGCAGAACGC CTCGGAGCGG GAGATCTTCT 1500
ACAAACGCCAC CGGCTTCCCT AGGAACCTCA GCTCAGCCTC TCAGGCCACT CGCCAGAAGA 1560
TGCGGGAGTG CCAGCGGCTG GTGACGCCCC TGCTCACCTC TATCAACCAC GCCCTGGAGC 1620
CGGGCAATG CAGGACAAG AGCGTGGAGA ACGCGTGTG GTCTCTGCGG AACCTGTCTT 1680
ACCGCTCTTA CGAGAGATG CCGCGSTCCG CGCTGCAGCG GCTGGAGGGT CGCGGCCGCA 1740
GGGACCTGGC GGGGGCGCG CCGGGAGAGG TCGTGGGCTG CTTACGCCCG CAGAGCCGCG 1800
GGCTGCGCGA GCTGCCCTTC GCGGCCGATG CGCTCACCTT CGCGGAGGTG TCCAAGSACC 1860
CCAAGGCCCT CGATGGCTG TGGAGCCCC AGATCGTGGG GCTGTACAAC CGGTGCTGTC 1920
AGCGCTCGCA GCTCAACCGG CACACGACGG AGCGGCCGCG CGGGGCGCTG CAGAACATCA 1980
CGGCAGGCGA CCGCAGTGG GCGGGGGTGC TGAGCCGCTT GGCCCTGGAG CAGGAGCGTA 2040
TTCTGAACCC CCTGTAGAC GTGTACAGGA CCGCCGACCA CCACAGCTG CGCTCACTGA 2100
CTGGCCTCAT CCGAAACCTG TCTCGGAACG CTAGGAACAA GGACGAGATG TCCACGAAGG 2160
TGGTGAGCCA CCTGATCGAG AAGCTGOCAG GCAGCGTGGG TGAGAAGTCG CCCCAGCCG 2220
AGGTGCTGCT CAACATCATA GCTGTGCTCA ACAACCTGGT GGTGGCCAGC CCCATCGCTG 2280
CCCGAGACCT GCTGTATTTT GACGACTCC GAAAGCTCAT CTTATCAAG AAGAAGCGGG 2340
ACAGCCCCGA CAGTGAGAG TCCTCCCGGG CAGCATCCAG CCTCTGGCC AACCTGTGGC 2400
AGTACACAA GCTCACCCGT GACTTTCGGG CGAAGGGCTA TCGGAAGGAG GACTTCTGG 2460
GCCATAGGT GAGCCCTTCT GGAGGAGAAG GTGACGTGGC CCAGGTCCA AGGACAGAC 2520
TCAGCTCCAG GCTGCTTGGC AGCCAGCCT GGAGGAGAAG GCTAATGAGG GAGGGGCCCC 2580
TCGCTGGGGC CCGTGTGTGC ATCTTTGAGG GTCTGSGGCC ACCAGGAGGG GCAGGGTCTT 2640
ATAGCTGGGG ACTTGCTTC CGCAGGGCAG GGGGTGGGGC AGGGCTCAAG GCTGCTCTGG 2700
TGATGGGGT GGTATGCCAG TCACATTGGC AGAGTGGGG GTTGCTGTG GCTTGGCAGT 2760
ATCTGGGAT AGCCAGCACT GGAATAAAG ATGGCCATGA ACAGTCACAA AAAAAAAAAA 2820
AAAAGGAATT C
```

Seq ID NO: 351 Protein sequence:
Protein Accession #: NP_009114.1

```
1      11      21      31      41      51
```

1 | 11 | 21 | 31 | 41 | 51 |
MQDGNPILSA IQPEAGVCSL ALPSDLQLDR RGAEGPEAER LRAARVQEQV RARLLQLGQQ 60
PRENGAEPSE PEAEATARGTS: RGQYHTLQAG FSSRSQGLSG DKTSGFRPIA KPAYSASPNS 120
SRSAVDLSCS RRLSSAHNGG SAPGAAGYGG AQPTTPMPTR PVSFHERGGV GSRADYDTLS 180
LRLSLRGPGG LDDRYSLVSE QLEPAATSTY RAFAYERQAS SSSSRAGGLD WPEATEVSPS 240
RTIRAPAVRT LQRFQSSHSR: RGVGGAVPGA VLEFVARAPS VRSLSLSLAD SGHLPDVHGF 300
NSYGSHTQL RLSSGFDDID: LPSAVKYLMA SDPNLQVLGA AYIQHKCYSD AAAKQARS 360
QAVPRLVLKF NHANQEVQRH ATGAMRNLIY DNADNKLALV EENGIFELLR TLREQDDEL 420
KNVTGILWNL SSSDHLKDR LARDTLEQLTD LVLSPLSGAG GPPLIQNAS EAEIFYNATG 480
FLRNLSSASQ ATRQKMRCH: GLVDALVTSI NHALDAGKCE DKSVENAVCV LRNLVSYRLD 540
EMPPSALQRL EGRGRDLAG APPGEVVGCP TPQSRRLREL PLAADALTFA EVSKDPKGL 600
WLWSPQIVGL YNRLQRCLE: NRHTTEAAAG ALQNTAGDR RWAGVLSRLA LEQERILNPL 660
LDRVRTADEH QLRSLTGLIR NLSRNARNKD EMSTKVVSHL IEKLPGSVGE KSPPAEVLVN 720
IIAVLNNLVV ASPIAARDLL YFDGLRKLIF IKKKRSDPDS EKSSRASSL LANLWQYNKL 780
HRDFRAKGYR KEDFLGP

Seq ID NO: 352 DNA sequence
Nucleic Acid Accession #: M31469
Coding sequence: 1-651

1 | 11 | 21 | 31 | 41 | 51 |
ATGGCTGCGC AGGGAGAGCC CCAGGTCCAG TTCAACTTG TATTGGTTGG TGATGGTGGT 60
ACTGGAAAAA CGACCTTCGT GAAACGTCAT TTGACTGGTG AATTGAGAA GAAGTATGTA 120
GCCACCTTGG GTGTGTAGGT TCATCCCTTA GTGTCCACA CCAACAGAGG ACCTATTAA 180
TTCATATGAT GGGACACAGC CGGCCAGGAG AAATTCGGTG GACTGAGAGA TGCTATTAT 240
ATCCAAGCCC AGTGTGCATC CATATGTTT GATGTACAT CGAGAGTTAC TTACAAGAA 300
GTGCTTAATC GGCATAGAGA TCTGTACGA GTGTGTGAAA ACATCCCCAT TGTGTGTGT 360
GGCAACAAG TGGATATTAA GGACAGGAAA GTGAAGCGCA AATCCATTGT CTTCACCGA 420
AAGAAGAAAT TTCACTACTA CGACATTTCT GCCAAAAGTA ACTACACTT TGAAAAGCCC 480
TTCTCTGGC TTGCTAGGAA GCTCATTTGA GACCTAACT TGGAAATTTG TGCCATGCCT 540
GCTCTGCCCC CACCAGAAT TGTCTAGGAC CCAGCTTTGG CAGCACAGTA TGACACGAC 600
TTAGAGGTTG CTACAGAAC TGCTCTCCCG GATGAGGATG ATGACCTGTG A

Seq ID NO: 353 Protein sequence
Protein Accession #: AAA36546

1 | 11 | 21 | 31 | 41 | 51 |
MAAQGEFQVQ FKLVLVDGG TGKTFVVRH LTGEFEKRYV ATLGVEVHPL VFHTNRGPIK 60
PNVWDTAGGE KFGGLRDGYI IQAQCALIMF DVTSRVTYRN VPNNHRLDVR VCENIPVL 120
GNKVDIKDRK VKAKSIVFHR KKNLQYDIS AKSNYNFEKP FLWLARKLIG DPNLEFVAMP 180
ALAPPEVVM D PALAAQYEH D LEVAQTALP DEDDDL

Seq ID NO: 354 DNA sequence
Nucleic Acid Accession #: NM_002820
Coding sequence: 304-831

1 | 11 | 21 | 31 | 41 | 51 |
CCGGTTCGCA AAGAAGCTGA CTTGAGAGG GGAACCTTC TTCTTTAGG AGGCGGTTAG 60
CCCTGTTCGA CGAACCCAGG AGAACTGCTG GCCAGATTAA TTGACATTG CTATGGGAGA 120
CGTGTAACA CACTACTTAT CATTGATGCA TATATAAAAC CATTTTATT TCGCTATTAT 180
TTGAGAGGAA GCGGCTCTGA TTTGTTCTT TTTTCCCTTT TTGCTCTTC TGCGTGTGTG 240
GTTTGGAGAA AGCACAGTTG GAGTAGCCGG TTGCTAAATA AGTCCCGAGC GCGAGCGGAG 300
ACGATGCAGC GGAGACTGGT TCAGCAGTGG AGCGTCCGGG TGTTCTGCTG GAGCTACGGC 360
GTGCCCTCCT GCGGCGCTCT GGTGGAGGGT CTCAGCCGCC GCCTCAAAAG AGCTGTGTCT 420
GAACATCAGC TCCTCCATGA CAAGGGGAAG TCCATCCAAG ATTTACGGCG ACGATTCTTC 480
CTTCAACATC TGATCGCAGA AATCCACACA GCTGAAATCA GAGCTACCTC GGAGGTGTCC 540
CCTAACTCCA AGCCCTCTCC CAACACAAAG AACCAACCCG TCCGATTGG GTCTGATGAT 600
GAGGCGAGAT ACCTAACTCA GGAACCTAAC AAGGTGGAGA CGTACAAAGA GCAGCCGCTC 660
AAGACACCTG GGAAGAAAAA GAAAGGCAAG CCGGGGAAC GCAAGGAGCA GGAAGAGAAA 720
AAACGGCGAA CTCGCTCTGC CTGCTTAGAC TCTGGAGTGA CTGGAGTGG GCTAGAGGG 780
GACCACCTGT CTGACACCTC CACAACTGCG CTGGAGCTCG ATTACCGGTA ACAGGCTTCT 840
CTGCGCCGTA GCCTCAGCGG GGTGCTCTCA GCTGGGTTT GGAGCCTCCC TTCTGCTTGT 900
GCTTGACAAA ACCTAGAATT TTCTCCCTTT ATGTATCTCT ATCGATTGTG TAGCAATTGA 960
CAGAGATAAA CTCAGATAT: TGTCTGCCTT AAAGCAGTAC CCCCCTACCA CACACACCCC 1020
TGCTCTCCAG CACCATAGAG AGGCGCTAGA GCCCATCTCT CTTTCTCCAC CGTACCCAA 1080
CATCAATCTT TTACCATCT ACCAAATAAT TTCATATTCA AGCTTCAGAA GCTAGTGACC 1140
ATCTTCATAA TTTGCTGGAG AAGTGATTTT CTTCCCTTCA CTCTCACACC TGGGCAAACT 1200
TTCTTCAGTG TTTTTCATTT CTTACGTTCT TTCACCTCAA GGGAGAAATAT AGAAGCAATT 1260
GATATTATCT ACAACACTG CAGAACAGCA TCATGTCTA AACGATTCTG AGCCATTCTC 1320
ACTTTTATT TAATTAATG TATTAAATTA AATCTCAAT: TTATTTAAT GTAAAGAACT 1380
TAAATATATG TTTAAACACA TGCCTTAAAT: TTGTTTAAT AAATTTAACT CTGGTTTCTA 1440
CCAGCTCATA CAAATAAAT GGTTCCTGAA AATGTTTAAG TATTAACCTA CAAGGATATA 1500
GGTTTTCTCT ATGTATCTTT TTGTTCTATT GCAAGATGAA ATAATTTTTC TAGGGTAATG 1560
CCGTAGGAAA AATAAACTT CACATTTAAA AAAAA

Seq ID NO: 355 Protein sequence
Protein Accession #: NM_002820

1 | 11 | 21 | 31 | 41 | 51 |
MQRRLVQWVS VAVFLLSYAV PSCGRSVEGL SRRLKRAVSE HQLLHDKGKS IQDLRRRFP 60
HLLIAEHTA BIRATSEVSP NSKPSPNTRN HPVRFSGDDE GRYLTTQETNK VETYKEQPLK 120

TPGKKKKGKP GKRKEQSKKK RRTRSAWLDS GVTGSGLEGD HLSDTSTTSL ELDSR

Seq ID NO: 356 DNA sequence

Nucleic Acid Accession #: NM 017522

Coding sequence: 1-2100

5

1 11 21 31 41 51
ATGGGCTCC CCGAGCCGG CCTCTCCGG CTCTGGGCG TGCTGCTGCT GCTGCTGCTG 60
CTGCTGCTGC TGGGCTCCA GCATCTTGG GCGGCGCGG CTGATCCGCT GCTCGGCGGC 120
CAAGGCGCCG CCAAGGAGTG CGAAAAGSAC CAATTCCAGT GCCGGAACGA GCGCTGCATC 180
CCCTCTGTGT GGAGATGCGA CGAGGAGGAT GACTGCTTAG ACCACAGCGA CGAGGACGAC 240
TGCCCCAAGA AGACCTGTGC AGACAGTGAC TTCACTGTGT ACAACGGCCA CTGCTATCCAC 300
GAACGCTGGG AGTGTGACGG CGAGGAGGAG TGCTCTGATG GCTCCGATGA GTCCGAGGCC 360
ACTTGCACCA AGCAGGTGTG TCCTGCAGAG AAGCTGAGCT GTGGACCCAC CAGCCACAAG 420
TGTTGACCTG CCTCTGTCGG CTGCGACGGG GAGAAGGACT GCGAGGGTGG AGCGGATGAG 480
GCGGCTGTGT CTACCTCACT GGGCACTGC CGTGGGACG AGTTCCAGTG TGGGGATGGG 540
ACATGTGTCC TTGCAATCAA GCATCTGCAAC CAGGAGCAGG ACTGTCCAGA TGGGAGTGAT 600
GAAGCTGGG GCGTACAGGG GCTGAACGAG TGCTGACACA ACAATGGCGG CTGCTCACAC 660
ATCTGCACTG ACCTCAAGAT TGGCTTTGAA TGCACTGCC CAGCAGGCTT CCAGCTCCTG 720
GACCAGAAGA CTGTGTCGGA CATTTGATGAG TGCAAGGACC CAGATGCCCTG CAGCCAGATC 780
TGTTGCAATT ACAAGGGCTA TTTTAAGTGT GAGTGTCTAC CTGGCTGCGA GATGGACCTA 840
CTGACCAAGA ACTGCAAGGC TGCTGTGCGG AAGAGCCCAT CCTTAATCTT CACCAACCGC 900
ACGAGTGGG AGGATCAGCG GTGAAGCGG AACTATTAC GCCTCATCCC CATGCTCAAG 960
AATGTCGTGG CACTAGATGT GGAAGTTGCC ACCAATCGCA TCTACTGGTG TGACCTCTCC 1020
TACCGTAAGA TCTATAGCCG CTACATGGAC AAGGCCAGTG ACCCGAAGA GCGGGAGGTC 1080
CTCATTGACG AGCAGTTGCA CTCTCCAGAG GCGCTGGCAG TGGACTGGGT CCACAAGCAC 1140
ATCTACTGGA CTGACTCGGG CAATAAGACC ATCTCAGTGG CCACAGTTGA TGGTGGCGCC 1200
GACCGCACTC TCTTACAGCG TAACCTCAGT GAACCCCGGG CCATCGCTGT TGACCCCTCTG 1260
CGAGGGTCCA TGTATTGGTC TGACTGGGGG GACCAGGCCA AGATTGAGAA ATCTGGGCTC 1320
AACGGTGTGG ACCGCGAAAC ACTGGTGTC GACAAATATT AATGGCCCAA CGGAATCACC 1380
CTGGATCTGC TGAGCCAGCG CTTGTACTGG GTAGACTCCA AGCTACACCA ACTGTCCAGC 1440
ATTGACTTCA GTGAGGGCAA CAGAAAGACG CTGATCTCCT CCATGACTTT CCTGAGCCAC 1500
CCTTTTGGGA TAGCTGTGTT TGAGGACAA GTGTTCTGGA CAGACCTGGA GAACGAGGCC 1560
ATTTTCAGTG CAAATCGGCT CAATGGCCTG GAAATCTCCA TCCTGGCTGA GAACCTCAAC 1620
AACCACATG ACATTGTCTAT CTTCATGAG CTGAAGCAGC CAAGAGCTCC AGATGCTCTG 1680
GAGCTGAGTG TCCAGCCTAA TGGAGGCTGT GAATACCTGT GCCTTCCTCG TCCTCAGATC 1740
TCCAGCCTAT CTCCCAAGTA CACATGTGCC TGCTCTGACA CAATGTGGCT GGGTCCAGAC 1800
ATGAAGAGGT GCTTCCGAGA TGCAAAATGA GACAGTAAGA TGGGCTCAAC AGTCACTGCC 1860
GCTGTTATCG GGATCATGCT GCCCATAGTG GTGATAGCCC TCCTGTGATC GAGTGGATAC 1920
CTGATCTGGA GAACTGGAA GCGGAAGAAC ACCAAAGACA TGAATTTTGA CAACCCAGTC 1980
TACAGGAAAA CAACAGAAGA AGAAGATGAA GATGAGCTCC ATATAGGGAG AACTGCTCAG 2040
ATTGGCCATG TCTATCTCTG ACGAGTGCCA TTAAGCCTTG AAGATGATGG ACTACCTCTG 2100
GGATGGGATC ACCGCCCTCG TGCTCTATGG AATTCACTCC CATGCACTAC ACTCCGGATG 2160
GTGATGACT GGATGAATGG GTTCTATAT ATGGGTCTGT GTGAGTGTAT GTGTGTGTGT 2220
GATTTTTTTT TTTAAATTTA TGTTCGGGAA AGGTAACCC AAAGTTATGA TGAATGCAA 2280
ACATCCAAAG GATGTGAGAG TTTTCTATG TATAATGTTT TATACACTTT TTAAGTGGTT 2340
GCACATCCCA TGAGGAATTC GTGGAATGGC TACTGCTGAC TAACATGATG CACATAACCA 2400
AATGGGGGCG AATGGCACAG TACCTTACTC ATCATTTAAA AACTATATT ACAGAAGATG 2460
TTTGGTGTCT GGGGGGCTTT TTTAGTITT GGGCATTGT TTTTGTAAA TAAGATGATT 2520
ATGCTTTGTG GCTATCCATC AACATAAGT

Seq ID NO: 357 Protein sequence

Protein Accession #: NP_059992

55

1 11 21 31 41 51
MGLPEPGLR LLALLLLLLL LLLRLQLHA AAAADPLLGG QGPAKECEKD QFQCRNERCI 60
PSVWRDDED DCLDHSDEDD CPKKTCAUSD FTDNGHCII BRWKCDGEEB CPDGSDESEA 120
TCTKQVPAE KLSGPTSHK CVPASWRCDG EKDCGGGADE AGCATSLGCT RGDEFQCGDG 180
TCVLAIKHCN QEQDCPDGSD EAGCLQGLNE CLHNNGGCSH ICTDLKIGFE CTCFAGFOLL 240
DQKTCGDIDE CQDPDACSQI CVNYKGYFKC ECYFGCEMDL LTNCKAAAG KSPSLIFTNR 300
TSAEDRPVKR NYSRLIPLMK NVVALDVEVA TNRIYWCDSL YRKIYSAYMD KASDPKEREV 360
LIDEQLHSPB GLAVDWHKHI IYWTDSGNKT ISVATVDGGR RRTLFNRNLS EPRAIADVPL 420
RGFMWSDWG DQAKIEKSLG NGVDRQTLVS DNIEWPNGIT LDLLSQRLYW VDSKLHQLSS 480
IDFSGNWRKT LISSTDFLSH PFGIAVFEDK VFWTDLENEA IPSANRLNGL EISILAENLN 540
NFHDIVIFHE LKQPRAPDAC ELSVQPNGGC EYLCLPAPQI SSHSPKTYCA CPDTMWLPGD 600
MERCYRDANE DSKMGSTVTA AVIGIIVFIV VIALLCMSGY LIWRNWKRNK TKSMNFDNPV 660
YRKTEBEDE DELHIGRTAQ IGHVYPARVA LSLEDGLP

70

Seq ID NO: 358 DNA sequence

Nucleic Acid Accession #: M27826

Coding sequence: <1-503

75

1 11 21 31 41 51
AGCCCAAGAA ACATCTCACC AATTTCAAAT CTGATCTATT GGGCTTAGCG ACTGAAGATT 60
GACGCTGCCC GATCGCTCGG GAAGTCCCTT GGACCATCAC AGAAGCCGAG CTTCGGGTAA 120
CTCTCAGAGT GGAGGGTAAG TCCATCCCTT GTTTAATCGA TACGGGGGCT ACCCACTCCA 180
CGTGTGCTTC TTTTCAAGGG CCTGTTTCCC TTGCCCCCAT AACTGTTGTG GGTATTGACG 240
GCCAAGCTTC AAAACCCCTG AAAACTCCCC CACTCTGGTG CCAACTTGGG CAACACTCTT 300
TTATGCACTC TTTTATTAGT ATCCCCACCT GCCCACTTCC CTATTATTAG CGAAATATTT 360
TAACCAAAAT ATCTGCTTCC CTGACTATTC CTGGAGTACA GCTACATCTC ATTGCTGCCC 420
TTCTTCCCAA TCCAAAGCCT CCTTTGTGTC CTCTAACATC CCCACAATAT CAGCCCTTAC 480
CACAAAGACT CCCCTCAGCT TAATCTCTCC CACTCTAGGT TCCACGCGG CCCCTAATCC 540
CACTTGAAGC AGCCCTGAGA AACATCGCCC ATTCTCTCTC CATACCACCC CCCAAAAATT 600
TTGCGCGCTC CAACACTTCA ACACATTTTT GTTTTATTGT TCTTATTAT ATCAGAAGGC 660

85

AGGAATGTCA GGCCTCTGAG CCCAGGCCAG GCCATGCGAT CCCTGTGAC TTGCAGGTAT 720
 ACATCCAGAT GGCCTGAAGT AACTGAAGAT CCACAAAAGA AGTAAAAACA GCCTTAACTG 780
 ATGACATTCC ACCATTGTGA TTTGTTCTGT CCCACCCCTA ACTGATCAAT GTACTTTGTA 840
 ATCTCCCCCA CCTTAAAGAA GGTTCCTTGT AATTCTCCCC ACCCTTGAGA ATGTACTTTG 900
 TGAGATCCAC CCTGCCCCAC CAGAGAACAA CCCCCTTTGA TTGTAATTTT TTATTACCTT 960
 CCAAAATCCT ATAAAAACAGC CCCACCCCTA TCTCTCTCA CTGACTCTCT TTTCGGACTC 1020
 AGCCACGGCG ACCCAGGTGA AATAAACAGC TTTATTGCTC AC

Seq ID NO: 359 Protein sequence
 Protein Accession #: AAA65999

1 11 21 31 41 51
 PKKHLTNFKS DLEGLATEDW RCPPIASEVPW TITEAELRVT LTVEGKSIPC LIDTGATHST 60
 LPSFQGPVSL APITVVGIDG QASKPLKTPP LWQQLQHSF MHSFLVIPTC PLPLLGRLN 120
 TKLSASLTIP: GVQLHLIALA LPNPKPLPCP LTSPQYQPLP QDLPSA

Seq ID NO: 360 DNA sequence
 Nucleic Acid Accession #: NM_001854
 Coding sequence: 162-5582

1 11 21 31 41 51
 AACCATCAAA TTTAGAAGAA AAAGCCCTTT GACTTTTTC CCCTCTCCCT CCCCAATGGC 60
 TGTGTAGCAA ACATCCCTGG CGATACCTTG GAAAGGACGA AGTGTGCTG CAGTCGCAAT 120
 TTCGTGGGTT GAGTTTCACAG TTGTGAGTGC GGGGCTCGGA GATGGAGCCG TGGTCTCTTA 180
 GGTGGAAGAC GAACGGGTGG CTCTGGGATT TCACCGTAAC AACCTCGCA TTGACCTTCC 240
 TCTTCCAAGC TAGAGAGGTC AGAGGAGCTG CTCAGTTGA TGTACTAAAA GCACTAGATT 300
 TTCACAATTC TCACAGAGGA ATATCAAAA CAACGGGATT TTGCACAAAC AGAAAGAATT 360
 CTAAGGGCTC AGATACTGCT TACAGAGTTT CAAGCAAGC ACAACTCAGT GCCCCAACAA 420
 AACAGTTATT TCCAGGTGGA ACTTCCCAG AAGACTTTTC AATACTATT ACAGTAAAC 480
 CAAAAAAGG AATTCAGTCT TCCCTTTTAT CTATATATAA TGAGCATGTT ATTACAGAAA 540
 TTGGTGTGGA GGTGGGAGAA TCACCTGTTT TCTGTTTGA AGACCACT GGAAGAACCTG 600
 TGGTGTGGA GGTGGGAGAA TCACCTGTTT TCTGTTTGA AGACCACT GGAAGAACCTG 660
 CCCCAGAGA: CTATCCCTCT TCCAGAACTG TTAACATCGC TGACGGGAAG TGGCATCGGG 720
 TAGCAATCAG: CGTGGAGAGG AAAAAGTGA CAATGATTGT TGATTGTAAG AAGAAACCA 780
 CGAAACCACT TGATAGAAGT GAGAGAGCAA TTGTGATAC CAATGGAATC ACGGTTTTTG 840
 GAACAAGGAT TTTGATGAA: GAAGTTTTTG AGGGGGACAT: TCACGAGTTT TTGATCAGAG 900
 GTGATCCCAA GGCAGCATAT GACTACTGTG AGCATTATAG TCCAGACTGT GACTCTTCAG 960
 CACCCAGAGC TGCTCAAGCT CAGGAACCTC AGATAGATGA GTATGACCA GAGGATATAA 1020
 TCGAATATGA CTATGAGTAT GGGGAAGCAG AGTATAAGA GGTGAAAGT GTAACAGAGG 1080
 GACCACTGT AACTGAGGAG: ACAATAGCAC AGACGGAGGC AAACATCGTT GATGATTTC 1140
 AAGAATACAA: CTATGGAACA ATGGAAGTT ACCAGACAGA AGCTCTTAGG CATGTTTCTG 1200
 GGCAAAATGA GCCAAATCAA GTTGAAGAAA TATTACTGA AGAATATCTA ACGGGAGAGG 1260
 ATTATGATT CAGAGGAAA AATTCTGAGG ATACACTATA TGAACAAAA GAAATAGACG 1320
 GCAGGGATTC TGATCTCTG GTAGATGGAG ATTTAGGCGA ATATGATTTC TATGAATATA 1380
 AAGAATATGA AGATAAACCA ACAAGCCCC CTAATGAAGA ATTTGGTCCA GGTGTACCAG 1440
 CAGAAACTGA TATTACAGAA ACAAGCATAA ATGGCCATGG TGCAATGGA GAGAAAGGAC 1500
 AGAAAGGAGA ACCAGCAGTG GTTGAGCTGT GTATGCTTGT CGAAGGACCA CAGGACCAAG 1560
 CAGGACCTGC AGGTATTATG GGTCTCCAG GTCTACAAG: CCCCCTGGA CCCCCTGGTG 1620
 ACCCTGGGCA TAGGGGCCCC CCAGGACGTC CTGGCTTACC AGGGGCTGAT GGTCTACCTG 1680
 GTCTCTCTGG TACTATGTTG ATGTTACCGT TCCGTTATGG TGGTGTGTT TCCAAAGGAC 1740
 CAACCATCTC TGCTCAGGAA GCTCAGGCTC AAGCTATTCT TCAGCAGGCT CGGATTGCTC 1800
 TGAGAGGCCC ACCTGGCCCA ATGGGTTCTA CTGGAAGACC AGGTCCTGTG GGGGGGCTC 1860
 GTTCATCTGG GCCCAAAGGT GAGAGTGTG ATCCAGTCC TCAGGGCCCT CGAGGCGTCC 1920
 AGGGTCCCC TGGTCCAAGC GGAAGACCTG GAAAAGGGG TCGTCCAGGT GCAGATGAGG 1980
 GAAGAGGAAT GCCAGGAGAA CCTGGGGCAA AGGGAGATCG AGGGTTTGTG GGAATCCCGG 2040
 GTCTGCGAGG TGACAAGGT CACAGGGGTG AACGAGGTCC TCAAGGTCTC CCAGGTCTCT 2100
 CTGGTGATGA TGAATGAGG CTGGGTCCAA GGGGAATCCC AGGAGTCCA GGGCAGCCTG 2160
 AAGCTGGCCC ACAGAGTTTG CTGGGTCCAA GGGGAATCCC AGGAGTCCA GGGCAGCCTG 2220
 GTATGCGAGG TGTAGATGCG CCCCCAGGAC CAAAAGGGAA CATGGGTCCC CAAGGGGAGC 2280
 CTGGGCTCC AGGTCAACAA GGGAAATCCAG GACCTCAGGG TCTTCTGTGT CCACAAGGTC 2340
 CAATTGGTCC TCTCTGTGAA AAGGAGCCAC AAGGAAAAAC AGGACTTGCT GGAATCTCTG 2400
 GTGCTGATG GCTCTCTGTT CATCTCTGGA AAGAAGGCCA GTCTGGAGAA AAGGGGCTC 2460
 TGGTCCCCC TGGTCCACAA GGTCTATTG GATNCCGGG CCCCCGGGA GTAAAGGGAG 2520
 CAGATGGTGT CAGAGGTCTC AAGGGATCTA AAGGTGAAA GGTGGAAGT GGTTTTCCAG 2580
 GATTCAAAGG TGACATGGGT CTAAAAGGTG ACAGAGGAGA AGTTGGTCAA ATTGGCCCAA 2640
 GAGGGAAGA TGGCCTGAA GGAACCAAG GTCCAGCAGG CCAACTGGA GACCCAGGTC 2700
 CTTGAGTCA AGCAGGAGAA AAGGGAACAC TTGGAGTTCC AGGATTACCA GGATATCCAG 2760
 GAAGACAAGG TCCAAAGGGT TCCACTGAT TCCCTGGGTT TCCAGGTGCC AATGGAGAGA 2820
 AAGGTGACAG GGGAGTAGCT GGCACACAG GCGCTCGGG TCAGCTGCTG CCAACGGGTC 2880
 CTCGAGGTTT AAGAGGTGCA AGAGGTCCCA CTGGGAAACC TGGGCCAAG GGCATTTCAG 2940
 GTGGGATGCG CCTCTCTGGC CCTCCAGGTG AAGAGGTCC TCAAGGACCT CAGGTCAGG 3000
 TTGGATTCCC TGGACCAAAA GGCCTCTCTG GACCACAGG AAGGATGGGC TGCCAGGAG 3060
 ACCCTGGGCA ACCTGGGGAG ACTGGATTTC AAGGCAAGAC CGGCCCTCCT GGGCAGGGG 3120
 GAGTGGTTGG ACCACAGGGA CCAACCGGTG AGACTGGTCC AATAGSGGAA CGTGGGTATC 3180
 CTGGTCTCTC TGGCCTCTCT GGTGAGCAAG GTCTCTCTGG TGCTGCAGGA AAAGAAGGTG 3240
 CAAAGGTGTA TCCAGTCTCT CAAGGTATCT CAGGGAAGA TGAACGACA GGATTACGTC 3300
 GTTTCACAGG GGAAGAGGT CTCTCTGGAG CTGAGGTGC ACCTGAGCTG AAAGGAGGGG 3360
 AAGGTCCCCA GGGCCACCA CGAGGGGCC CCGGACCTCA GGGTCTCTCT GGTCCAGCTG 3420
 CAGCTGGCCC AATTGGTTTA CGAGGGGCC CCGGACCTCA GGGTCTCTCT GGTCCAGCTG 3480
 GAGAGAAAG TGCTCTGGA GAAAAGGTC CCAAGGGCC TCAGGGAGA GATGGAGTTT 3540
 AAGGTCTCTG TGGTCTCCCA GGGCCAGCTG GTCTCTCGGG CTCCCTGGG GAAGACGGAG 3600
 ACAAGGGTGA AATTGGTGA CCGGACAAA AAGGCAGCAA GGTGGCAAG GGAGAAATG 3660
 GCGCTCCCCG TCCCCAGGT CTTCAAGGAC CAGTGGTGC CCTTGAATT GCTGGAGGTG 3720
 ATGGTGAACC AGGTCTCTGA GGACAGCAGG GGATGTTTGG GCAAAAAGGT GATGAGGGTG 3780
 CCAGAGGCTT CCTGGACCT OCTGGTCCA TGGTCTTCA GGGTCTGCA GGGCCACTG 3840
 GTGAAAAGG TGAATAAGG GATGTTGTC CATGGGGGCC ACCTGCTCT CAGGCCCCAA 3840

5
10
15
20
25
30
35
40

GAGGCCCTCA AGGTCCCAAT GGAGCTGATG GACCACAAGG ACCCCAGGT TCTGTTGGTT 3900
CAGTTGGTGG TGTGGAGAA AAGGGTGAAC CTGGAGAAGC AGGAAACCCA GGGCTCTCTG 3960
GGGAAGCAGG TGTAGGCGGT CCCAAAGGAG AAAGAGGAGA GAAAGGGGAA GCTGGTCCAC 4020
CTGGAGCTGC TGGACCTCCA GGTGCCAAGS GGGCGCCAGG TGATGATGGC CCTAAGGGTA 4080
ACCCGGGTCC TGTGTTGTTT GACAAGGGTG AAGATGGAGA TCCTGGTCAA CGGGTCTCTC 4140
GTCAAGATGG TGTGTTGTTT GAGCAAGGGT AAGATGGAGA TCCTGGTCAA CGGGTCTCTC 4200
CTGGCCCATC TGGTGAAGCT GGGCCACCAG GTCTCTCTGG AAAACGAGGT CCTCTGGAG 4260
CTGACGGTGC AGAGGCAAGA CAAGGTGAAA AAGGTGCTAA GGGGGAAGCA GGTGCAGAA 4320
GTCTCTCTGG AAAACCGGCT CCAGTGGGTC CTCAGGGACC TGCAGGAAGG CCTGGTCCAG 4380
AAGGTCTCTG GGGCATCCCT GGTCTCTGGG GAGAACAAGS TCTCCCTGGA GCTGCAGGCC 4440
AAGATGGACC ACCTGGTCTC ATGGGACCTC CTGGCTTACC TGGTCTCAA GGTGACCTG 4500
GCTCCAGGGT TGAAGAGGGA CATCTCTGTT TAATTGGCCT GATTGGTCTC CCAGGAGAAC 4560
AAGGGGAAA AGGTGACCGA GGGCTCCCTG GAACTCAAGG ATCTCCAGGA GCAAAAGGGG 4620
ATGGGGAAT TCCTGGTCTC GCTGGTCCCT TAGTCCACC TGGTCTCCA GGCTTACCAG 4680
GTCTCTCAAG CCCAAAGGGT AACAAAGGCT CTACTGGACC CGCTGGCCAG AAAGGTGACA 4740
GTGGTCTTCC AGGGCTCTCT GGGCTCCAG GTCCACCTGG TGAAGTCATT CAGCTTTAC 4800
CAATCTTGTG CTCACAAAAT ACGAGAAGAC ATACTGAAGS CATGCAAGCA GATGCAGATG 4860
ATAATATTCT TGATTACTCG GATGGAATGG AAGAAATATT TGGTTCCTCC AATTCCCTGA 4920
AACAGACAT CGAGCATATG AAATTTCCAA TGGGTACTCA GACCAATCCA GCCCGAAGCT 4980
GTAAAGACCT GCAACTCAGC CATCTGACT TCCAGATGG TGAATATTGG ATTGATCTTA 5040
ACCAAGTTTG CTCAGGAGAT TCCTTCAAAG TTTACTGTAA TTTACATCT GGTGGTGA 5100
CTTGCAATTA TCCAGACAAA AAATCTGAGG GAGTAAGAAT TTCATCATGG CCAAGGAGA 5160
AACCGAAGT TGTGTTTAGT GAATTTAAGA GGGGAAAAC GCTTTCATAC TTAGATGTTG 5220
AAGGAAATTC CATCAATATG GTGCAATGA CATCTCTGAA ACTTCTGACT GCCTCTGCTC 5280
GGCAAAATTT CACCTACCAC TGTCTACAGT CAGCAGCCTG GTATGATGTG TCATCAGGAA 5340
GTTATGACAA AGCACTTCGC TTCTGGGAT CAAATGATGA GGAGATGCTC TATGACAATA 5400
ATCTTTTAT CAAAACACTG TATGATGTTT GTACGTCCAG AAAAGGCTAT GAAAAAACTG 5460
TCATTGAAAT CAATACACCA AAAATTGATC AAGTACCTAT TGTGATGTC ATGATCAGTG 5520
ACTTTGGTGA TCAGAACTAG AAGTTCGGAT TTGAAGTTGG TCCTGTTTGT TTTCTTGGCT 5580
AAGATTAGA CAAAGAACAT ATCAATCAA CAGAAATGT ACCTTGGTGC CACCAACCCA 5640
TTTTGGCCA CATGCAAGTT TTGAATAAGG ATGTATGGAA AACACGCTG CATATACAGG 5700
TACCATTAG GAAATACCGA TGCCCTTGTG GGGGAGCAAT CACAGACAAA AGCTTTGAAA 5760
ATCATAAAGA TATAAGTTGG TGTGGCTAAG ATGGAAACAG GGCTGATTCT TGATTCCTAA 5820
TTCTCAACTC TCCTTTTCTC ATTTGAATTT CTTTGGTGTG GTAGAAAAA AAAAAAGAAA 5880
AATATATATT CATAAAAAAT ATGGTGCTCA TTCTCATCCA TCCAGGATGT ACTAAACAG 5940
TGTGTTAAT AAATTTGTAAT TATTTTGTG ACAGTTCTAT ACTGTTATCT GTGTCCATT 6000
CCAAACTTG CAGCTGTCCC TGAATCCGC TGACTCTAAT TTATGAGGAT GCCGAAGCT 6060
GATGGCAATA ATATATGTAT TATGAAATG AAGTTATGAT TTCCGATGAC CCTAAGTCCC 6120
TTTCTTGGT TAATGATGAA ATCTCTTGT GTGTGTTT

Seq ID NO: 361 Protein sequence
Protein Accession #: NP_001845

45
50
55
60
65
70
75

1 11 21 31 41 51
MEPWSSRWKT KRWLWDFTVT TLALTFLLQA REVRGAAPVD VLKALDFHNS PEGISKTTGF 60
CTNRKNSKGS DTAAYRVSKQA QLSAPTQQLF PGGTFPPEDFS ILFTVKKPKG IQSFLLSIYN 120
EHGLQIQIVE VGRSPVLFVE DHTGKPAFED YPLFRVTNIA DGKWHRAVIS VEKRTVTMIV 180
DCKKKTTRPL DRSERAIVDT NGITVFETRI LDEEVFEGDI QQFLITGDPK AAYDYCEHYS 240
PDCDSSAPKA AQAQEPQIDE YAPEDIIEYD VEYGEAEYKE AESVTEGPTV TEETIAQTEA 300
NIVDDPQEN YGTMBEYQTE APRHVSQTNE PNPVEISFTE EYLITGEDYDS QRKNSDITLY 360
ENKEIDGRDS DLLVDGLDGE YDFEYKEYE DKPTSPPNEE FGPGVPAETD ITETSINGHG 420
AYGEKQKGE PAVVEPGMLV EGPPGPAGPA GIMGPPGLQG PTGPPGDPGD RGPGRPGLP 480
GADGLPGPG TMLMLPFRYG GDGSKGPTIS AQEAQAQAIL QQARIALRGP PGPMGLTRP 540
GPDGLPGLPG DKHGRGERGP QGPPGPPGDD GMRGEDGEIG PRGLPGEAGP RGLLGPRTG 600
GAPGQPGMAG VDGPPGPKGN MGPPQGEPPG QQQGNPFPQG LPGPQGPFGP PGEKGPQKGP 660
GLAGLPGADG PPGHPKEGGQ SGEKALGFP GPQGPFGXPG PRGVKAGDGV RGLKSGKEK 720
GEDGPPGFKG DMGLKGRDGE VQIGPRGXD GPEGPKGRAG PTGDFPSPQG AGEKGLGVP 780
GLPGYPPRQG PKGSTGPPGF PGANGKCAR GVAGKPGPRG QRGPTGPRGS RGARGPTGKP 840
GPKGTSGDQ PPGPPGERGP QGPQGPVGF PPKPPGPPG RMCGPCPHPG RGTGFGQKT 900
GPPGPGVVG PGGTGTGTGP IGERGYPPGP GPPGQGLPG AAGKEGAGD PGPQGISGKD 960
GPAGLRGPPG ERGLPCAQA PGLKGGEGFP GPPGPVSPG ERGSAGTAGP IGLRGRPGP 1020
GPPGPAGEKG APGEKPGQGP AGRDGVQGV GLPGPAGPAG SPGEDGDKGE IGPQPKGSK 1080
GGKGENPPG PPLGQPVGA PGLAGDGER GPRGQQGMFG QKGDGARGF PGPPGPIGLQ 1140
GLPGPPGEKG ENGDVGPWG PGPPGPRGP GPNGADGPG PPGSVGVSG VGEKGEPEGA 1200
GNPGPPEAG VGGPKGERGE KGEAGPPGAA GPPGAKGPP DDGPKGNPFP VGFPDGPFP 1260
GELGPAQDQ VGGDKGEDG PQQPPGPPS GEAGPPGPPG KRGPAGAAGA EGRQGEKGA 1320
GEAGAEPPG KTGVPVQGP AGKPPGEGRL GIPGVPGEQG LPGAAQDGP PGPMGPPGLP 1380
GLKGDGSKG ERHGPGLGL IGPPGEQGEK GDRGLPGTQG SPGAKGDDGI PGFAGPLGPP 1440
GPPGLPGPQG PKGNKGSTGP AGQKGDGSLP GPPGPPGPPG EVIQPLPLS SKKTRRHTG 1500
MQADADDNIL DYSDEEIEF GSLNSLKQDI EHMKFPMTQ TNPARTCKDL QLSHPDFPDG 1560
EYWDPNQGC SDFSFKVYCN FTSGGETCIY PDKKSEGVRI SSWPKERKPS WFEFKRGLK 1620
LSYLDVEGNS INMVQMTFLK LLTASARQNF TYHQSAAM YDVSSGSYDK ALRFLGSNDE 1680
EMSYDNNPFI KTLYDGCISR KGYEKTVEI NTPKIDQVPI VDMISDFGD QNKFGFEVG 1740
PVCPLG

Seq ID NO: 362 DNA sequence
Nucleic Acid Accession #: NM_003107
Coding sequence: 351-1775

85

1 11 21 31 41 51
TTCCCCAGCA TTCCAGAAAC TCCTCTCTAC TTTAGCACGG TCTCCAGACT CAGCCGAGAG 60
ACAGCAAACT GCAGGCGCGT GAGAGAGCGA GAGAGAGGGA GAGAGAGACT CTCCAGCCTG 120
GGAACATAA CTCCTCTCGC AGAGGCGGAG AACTCCTTCC CCAATCTTT TGGGGACTTT 180

	TCTCTCTTA	CCCACCTCCG	CCCCTGCGAG	GAGTTGAGGG	GCCAGTTCGG	CCGCGCGCGG	240
	GGTCTTCCCG	TTGCGGTGT	GCTTGGCCCG	GGGAACCGGG	AGGGCCCGGC	GATGCGCGGG	300
	CGGCGCGCGG	GAGGGTGTGA	GCGGCGGTGG	GCGCCCGCGG	AGCGGAGGCC	ATGGTGCGAG	360
5	AAACCAACAA	TGCGAGAAC	ACGGAAGGCG	TGCTGGCCGG	CGAGAGCTCG	GACTCGGGCG	420
	CCGGCCTCGA	GCTGGGAATC	GCCTCCTCCC	CCAAGCCCGG	CTCCAACCGC	TCCACGGGGG	480
	GCAAGGCCGA	CGAAGCGAGC	TGGTGCAAGA	CCCCGAGTGG	GCACATCAAG	CGACCCATGA	540
	ACGCCCTTCAT	GGTGTGTGCG	CAGATCGAGC	GGCGCAAGAT	CATGGAGCAG	TGCGCCGACA	600
	TGCACAACGC	CGAGATCTCC	AAGCGGCTGG	GCAACGCTG	GAAAGTCTCT	AAAGACAGCG	660
10	ACAGATCCC	TTTCATTGGA	GAGGCGGAGC	GGCTGGCGCT	CAAGCACATG	GCTGACTACC	720
	CCGACTACAA	GTACCGGCCC	AGGAAGAAGG	TGAAGTCCGG	CAACGCCAAC	TCCAGCTCCT	780
	CGGCGCGCGC	CTCTCCCAAG	CCGGGGGAGA	AGGGAGACAA	GGTGGGTGGC	AGTGGCGGGG	840
	GCGGCCATGG	GGGCGCGCGG	GGCGGCGGGA	GCAGCAACGC	GGGGGGAGGA	GGCGGCGGTG	900
	CGAGTGGCGG	CGGCGCAAC	TCCAAACCGG	CGCAGAAAAA	GAGCTGCGGC	TCCAAAGTGG	960
15	CGGCGGCGCG	GGGCGGTGGG	GTTAGCAAA	CGCAGGCCAA	GCTCATCTG	CGAGGCGGGG	1020
	GCGGCGGCGG	GAAAGCAGCG	GCTGCCGCGG	CCGCTCCTT	CGCCCGCGAA	CAGGCGGGGG	1080
	CGGCGGCGCT	GCTGCCCTTG	GGCGCGCGCG	CCGACCACCA	CTCGGTGTAC	AAGGCGCGGA	1140
	CTCCAGCGCG	CTCGCGCGCG	CGCTCCTCGG	CAGCCTCGGC	CTCGGACAGC	CTCGCGGGCG	1200
	CGGCGAAGCA	CCTGGGCGAG	AAGAAGGTGA	AGCGGCTCTA	CCTGTTGCGC	GGCCTGGGCA	1260
20	CGTGTGCTG	GCGCGGTGGG	GGCGTGGGCG	CGGGAGCGGA	CCCCAGCGAC	CCCTGGGGCG	1320
	TGTACGAGGA	GGAGGCGCGG	GGCTGCTCGC	CCGACGCGCC	CAGCCTGAGC	GGCGCGAGCA	1380
	GCGCGGCGCT	GTCCCGCGCG	GCGGCGCGCT	CGCCCGCGGA	CCACCGCGGC	TACGCCAGCC	1440
	TGGCGCGCGC	CTCGCGCGCG	CGGTCCAGCG	CGCCCTCGCA	CGCGTCTCTC	TGCGCGCTCGT	1500
	CCCACTCCTC	CTCTCTCTCC	TCTTCGGGCT	CCTCGTCTCT	CGACGACGAG	TTGCAAGACG	1560
25	ACCTGCTCGA	CCCGGTGGCG	AGCTCAAACT	TTGAGAGCAT	GTCCTTGGGC	AGCTTCAGTT	1620
	CGTGTGCGG	GCTCGACCGG	GACCTGGATT	TAACTTCTGA	GCCCGGCTCC	GGCTCGCACT	1680
	TGAGTTTCCC	GGACTACTCG	ACGCGCGAGG	TGAGCGAGAT	GATCTCGGGA	GACTGGCTCG	1740
	AGTCCAGCAT	CTCCAACTTG	GTTTTCACCT	ACTGAAGGGC	CGCGAGGCGG	GGAGAAAGGG	1800
	CGGGGGGGGT	AGGAGAGGAG	AAAAAAAGAG	TGAAAAAAG	AAACGAAAG	GACAGACGAA	1860
30	GAGTTTAAAG	AGAAAGGGA	AAAAAGTAAG	AAAGAGTAAG	CAGGGCTCGT	TCCCGCGCGT	1920
	TCTCGTCTG	GGATCAAGGA	GCGCGCGCGC	GTTTGGAGCC	CGCGCTCCCA	TCCCGCACCT	1980
	TCCCGGGCGG	GGGAGCCACT	CTGCCAGCGC	GGAGGGAGCG	GGAGGAGGAA	GAGGGTAGAC	2040
	AGGGGGGAGC	TGTGATTGTT	GTTATTGATG	TTGTTGTTGA	TGGCAAAAAA	AAAAAGCGAC	2100
	TTGAGATTGG	CTCCCTCTTG	CTTGAAGAGA	CCCCCTCCCC	CTTCCAACGA	GCTTCCGGAC	2160
35	TTGTCTGCAC	CCCCAGCAAG	AAGGCGAGTT	AGTTTCTTAG	AGACTTGAAG	GAGTCTCCCC	2220
	CTTCTGTCAT	CACCACCTTG	GTTTTGTTTT	ATTTTGCTTC	TTGGTCAAGA	AAGGAGGGGA	2280
	GAACCCAGCG	CACCCCTCCC	CCCTTTTTTT	TAAACGCGTG	ATGAAGACAG	AAGGCTCCGG	2340
	GGTGACGAAT	TTGGCCGATG	GCAGATGTTT	TGGGGGAACG	CCGGGACTGA	GAGACTCCAC	2400
	GCAGGCGAAT	TCCGTTTGGG	GGCCTTTTTT	TCTTCCCTCT	TTTCCCTCTG	CCCCCTCTGC	2460
40	AGCCGAGGGA	GGAGATGTTG	AGGGGAGGAG	GCCAGCCAGT	GTGACCGGCG	CTAGGAAATG	2520
	ACCCGAGAAC	CCCGTTGGAA	GCGCAGCAGC	GGGAGCTAGG	GCGCGGGGCG	GAGGAGGACA	2580
	CGAACTGGAA	GGGGGTTCCG	GGTCAAACTG	AAATGGATTG	GCACGTGGGG	GAGCTGGGGG	2640
	CGGCGGCTGC	TGGGCTCCGC	CCTTCTTTTC	TACGTGAAAT	CAGTGAGGTG	AGACTTCCCA	2700
	GACCCCGGAG	GGGTGGAGGA	GAGGAGACTG	TTTGATGTGG	TACAGGGGCA	GTCAGTGGAG	2760
45	GGCGAGTGGT	TTCGGAAAAA	AAAAAGAAA	AAAAGGG			

Seq ID NO: 363 Protein sequence
Protein Accession #: NP_003098

	1	11	21	31	41	51	
50	MVQQTNNNAEN	TEALLAGESS	DSGAGLELGI	ASSPTPGSTA	STGGKADDPG	WCKTPSGHIK	60
	RPMNAFMVWS	QIERRKIMEQ	SPDMHNAEIS	KRLGRKWLKL	KDSKIPFIR	EAERLRLKHM	120
	ADYDPYKVRP	RKVKVSGNAN	SSSSAAASSK	PGEKGDVKVG	SGGGHGGGGG	GGSSSNAGGG	180
	GGGASGGGAN	SKPAKKSQCG	SKVAGGAGGG	VSKPHAKLIL	AGGGGGGKAA	AAAAASFPAE	240
55	QAGAAALLPL	GAAADHHSLY	KARTPSASAS	ASSAASASAA	LAAPGRHLAE	KKVKRVYLFQ	300
	GLGTSSSPVG	GVGAGADPSD	PLGLYEEBGA	GCSPDAPSLG	GRSSAASSPA	AGRSPADHRG	360
	YASLRAASPA	PSSAPSHASS	SASSHSSSSS	SSGSSSSDDG	FEDDLLDLNP	SSNFESMSLG	420
	SPSSSSALDR	DLDFNPEPGS	GSHFEFPDYC	TPEVSEMISG	DWLESSISNL	VFTY	

Seq ID NO: 364 DNA sequence
Nucleic Acid Accession #: U10860
Coding sequence: 123-2204

	1	11	21	31	41	51	
65	TGCGGGCTGC	TCTCGACCA	GGCCTCCTTC	TCAACCTCAG	CCGCGGCGGC	CGACCCCTCC	60
	GGCACCCCTCC	CGCCCGCTCT	CGTACTGTGG	CCGTACCGCG	CGCGGCTCCG	GGCCTGGCCC	120
	CGATGGCTCT	GTGCAACGGA	GACTCCAAGC	TGGAGAATGC	TGGAGGAGAC	CTTAAGGATG	180
70	GCCACCACCA	CTATGAAGGA	GCTGTGTGCA	TTCTGGATGC	TGGTGCTCAG	TACGGGAAAG	240
	TGATAGACCG	AAGAGTGAGG	GAATGTTTCG	TGCAGTCTGA	AATTTTCCCC	TTGGAAACAC	300
	CAGCATTTTC	TATAAAGGAA	CAAGGATTCC	GTGCTATTAT	CATCTCTGGA	GGACCTAATT	360
	CTGTGTATGC	TGAAGATGCT	CCCTGGTTTG	ATCCAGCAAT	ATTCACTATT	GGCAAGCCTG	420
75	TTCTTGAAT	TTGCTATGGT	ATGCAGATGA	TGAATAAGGT	ATTGGAGGT	ACTGTGCACA	480
	AAAAAAGTGT	CAGAGAAGAT	GGAGTTTTC	ACATTAGTGT	GGATAATACA	TGTTCAATTAT	540
	TGAGGGGCTC	TGAGAAGGAA	GAAGTTGTTT	TGCTTACACA	TGGAGATAGT	GTAGACAAAG	600
	TAGCTGATGG	ATTCAAGGTT	GTGGCAGGTT	CTGGAACAT	AGTAGCAGGC	ATAGCAAAAG	660
	AATCTAATAA	GTTATATGGA	GCACAGTTCC	ACCCTGAAGT	TGGCCTTACA	GAAATGGAA	720
80	GAATAATCT	GAAAGATTTC	CTTTATGATA	TAGCTGGATG	CAGTGGAAAC	TTACCGGTGC	780
	AGAACAGAGA	ACTTGAGTGT	ATTCGAGAGA	TCAAAGAGAG	AGTAGGCACG	TCAAAGTTT	840
	TGGTTTACT	CAGTGGTGGG	GTAGACTCAA	CAGTTTGTAC	AGCTTTGCTA	AATCGTGCTT	900
	TGAACCAAGA	CAAGGTCAAT	GCTGTGCACA	TTGATAATGG	CTTTATGAGA	AAACGAGAAA	960
	GCCAGTCTGT	TGAAGAGGCC	CTCAAAAGC	TTGGAATTCA	GGTCAAAGTG	ATAAATGCTG	1020
85	CTCATTCTTT	CTACATGGA	ACAAACACCC	TACCAATATC	AGATGAAGAT	AGAACCCAC	1080
	GGAAAAAGAT	TAGCAAAACG	TTAAATATGA	CCACAAGTCC	TGAAGAGAAA	AGAAAAATCA	1140
	TTGGGGATAC	TTTGTGTAAG	ATTGCCAATG	AAGTAATTGG	AGAAATGAAC	TTGAAACCAG	1200
	AGGAGGTTTT	CCTTGCCCAA	GGTACTTTAC	GGCCTGATCT	AATTGAAAGT	GCATCCCTTG	1260

5 TTGCAAGTGG CAAAGCTGAA CTCATCAAAA CCCATCACAA TGACACAGAG CTCATCAGAA 1320
 AGTTGAGAGA: GGAGGGGAAA GTAATAGAAC CTCTGAAAAG TTTTCATAAA GATGAAGTGA 1380
 GAATTTTGGG CAGAGAAGCTT GGAGCTCCAG AAGAGTTAGT TCCAGGAGAT CCATTTCAG 1440
 GTCTGGGCTT GGCAATCAGA GTAATATGTG CTGAAGAACC TTATATTGTG AAGGACTTTC 1500
 CTGAAACCAA CAATATTTTG AAAATAGTAG CTGATTTTTC TGCAAGTGTG AAAAGGCCAC 1560
 ATACCTTATT ACAGAGAGTC AAAGCCTGCA CAACAGAAGA GGATCAGGAG AAGCTGATGC 1620
 AAATTACCAAG TCTGCATTCG CTGAATGCCT TCTTGCTGCC AATTAAAACT GTAGGTGTGC 1680
 AGGGTGAAGT TCGTCTCTAC AGTTACGTGT GTGGAAATCTC CAGTAAAGAT GAACCTGACT 1740
 10 GGGAACTACT TATTTTCTG GCTAGGCTTA TACCTCGCAT GTGTCAACAAC GTTAACAGAG 1800
 TTGTTTATAT ATTTGGCCCA CCAGTTAAAG AACCTCTAC AGATGTTACT CCCACTTCT 1860
 TGACAAACAG GGTGCTCAGT ACTTTAGGCC AAGCTGATTT TGAGGCCCAT AACATTCTCA 1920
 GGGAGTCTGG GTATGCTGGG AAAATCAGCC AGATGCCGGT GATTTTGACA CCATTACATT 1980
 TTGATCGGGA CCCACTTCAA AAGCAGCCTT CATGCCAGAG ATCTGTGGTT ATTCGAACCT 2040
 15 TTATTACTAG TGACTTCATG ACTGGTATAC CTGCAACACC TGGCAATGAG ATCCCTGTAG 2100
 AGGTGTATT AAAGATGGTC ACTGAGATTA AGAAGATTCC TGGTATTCTT CGAATTATGT 2160
 ATGACTTAAC ATCAAGGCC CCAGGAACCTA CTGAGTGGGA GTAATAAAT TC

Seq ID NO: 365 Protein sequence

Protein Accession #: AAA60331

20 1 11 21 31 41 51
 MALCNGDSKL ENAGGDLKDG HHHYEGAVVI LDAGAQQYKV IDRRVRELFV QSEIFPLETP 60
 AFAIKQGFPR AIIISGGPNS VYAEADAPWF PAIFTIGKPV LGICYGMQMM NKVFGGTVEK 120
 25 KSVREDGVFN ISVDNCSLF RGLQKEEVL LTHGDSVDKV ADGKVVARS GNIVAGIANE 180
 SKKLYGAQFH PEVGLTENGL VILKNFLYDI AGCSGTFVQ NRELECIREI KERVGTSKVL 240
 VLLSGGVDS TCTALLNRL NQEQVIAVHI DNGFMKRRES QSVEEALKKL GIQVKVINA 300
 HSPYNGTTL PISDEBTRPR KRISKTLNMT TSPEEKRII GDTFVKIANE VIGEMNLKPE 360
 EVFLAQGTLR PDIIESASLV ASGKAEIKT HNDTELIRK LREEGKVIEP LKDFHKDEVR 420
 30 ILGRELGLPE ELVSRHPFPF PGLAIRVICA EEPYICKDFP ETNNILKIVA DFSASVKKPH 480
 TLLQVRKACT TEEDQEKLMO ITSLSLNAF LLPIKTVGVQ GDCRSYSYVC GISSKDEPDW 540
 ESLIFLARLI PRMCHNVNRV VYIFGPPVKE PPTDVTPTFL TTGVLSTLRQ ADFAHNILR 600
 ESGYAGKISQ MPVILTPLHF DRDPLQKQPS QRSVVIRTF ITSDFMTGIP ATPGNEIPVE 660
 35 VVLKMTVEIK KIPGISRIMY DLTSKPPGTT EWE

Seq ID NO: 366 DNA sequence

Nucleic Acid Accession #: NM_004219

Coding sequence: 46-654

40 1 11 21 31 41 51
 GGGGCTCAG ATGAATGCGG CTGTTAAGAC CTGCAATAAT CCAGAATGGC TACTCTGATC 60
 TATGTTGATA AGGAAATGG AGAACCAAGC ACCCGTGTGG TTGCTAAGGA TGGGCTGAAG 120
 45 CTGGGCTCTG GACCTTCAAT CAAAGCCTTA GATGGGAGAT CTCAAGTTTC AACACCAAGT 180
 TTGGCAAAA CGTTCGATGC CCCACCAGCC TTACCTAAGC CTACTAGAAA GGCTTTGGGA 240
 ACTGTCAACA GAGCTACAGA AAGTCTGTGA AAGACCAAGG GACCCCTCAA ACAAAAACAG 300
 CCAAGCTTTT CTGCCAAAAA GATGACTGAG AAGACTGTGA AAGCAAAAAG CTCTGTCTCT 360
 GCTCAGATG ATGCCATATC AGAAATAGAA AAATCTTTC CTTCTCAATCC TCTAGACTTT 420
 50 GAGAGTTTGT ACCCTGCTGA AGAGCACCAG ATTGGGCACC TCCCTCTGAG TGGAGTGCCT 480
 CTCATGATCC TTGACGAGGA GAGAGAGCTT GAAAAGCTGT TTCAGCTGGG CCCCCCTTCA 540
 CCTGTGAAGA TGCCCTCTCC ACCATGGGAA TCCAATCTGT TGCACTCTCC TTCAAGCAT 600
 CTGTGACACC TGGATGTTGA ATTGCCACT GTTGTCTGTG ACATAGATAT TTAATTTCT 660
 55 TAGTGCTTCA GAGTTTGTGT GTATTGTAT TAATAAGCA TTCTTCAACA GAAAAAAA 720
 AAAAAA

Seq ID NO: 367 Protein sequence

Protein Accession #: NP_004210

60 1 11 21 31 41 51
 MATLIYVDKE NGEPGTRVVA KDGLKLGSGP SIKALDGRSQ VSTPRFGKTF DAPPALPKAT 60
 RKALGTVNRA TEKSVKTRGP LKQKQSFSA KMTKTVKA KSSVPASDDA YPEIEKFFFP 120
 65 NPLDFESFDL PEEHQIAHLP LSGVPLMILD EERELEKLFQ LGPPSPVRMP SPWESNLLQ 180
 SPSSILSTLD VELPPVCCDI DI

Seq ID NO: 368 DNA sequence

Nucleic Acid Accession #: NM_000597

Coding sequence: 118-1104

70 1 11 21 31 41 51
 ATTCGGGGCG AGGGAGGAGG AAGAAGCGGA GGAGGCGGCT CCGCTCGCA GGGCGGTGCA 60
 75 CCGCCCCGCC CGCCCGCTCG CTCGCTCGCC CGCCGCGCCG CGCTGCGAC CGCCAGCATG 120
 CTGCGAGAGG TGGGCTGCCC CGCGCTGCGG CTGCGCGCGC CGCCGCTGCT GCGCTGCTG 180
 CGCTGCTGCT TGTCTCTACT GGGCGCGAGT GCGCGCGCGG GCGGGGCGCG CGCGGAGGTG 240
 CTGTTCCGCT GCCCGCCCTG CACACCCGAG CGCCTGGCGG CCGTGGGGCC CCGCGCGGTT 300
 GCGCGCGCGG CGCGGCTGCG CGCAGTGGCC GGAGGCGCGG GCATGCCATG CGCGGAGCTC 360
 80 GTCCGGGAGC CGGCTGCGG CTGCTGCTCG GTGTGCGCCC GGCTGGAGGG CGAGGCGTGC 420
 GGCGTCTACA CCGCGGCTG CGGCCAGGGG CTGCGCTGCT ATCCCCACCC GGGCTCCGAG 480
 CTGCCCCCTG AGCGCTGCTT CATGGGCGAG GGCATTTGTG AGAAGCGCGG GGACGCCGAG 540
 TATGCGGCGA CGCCGGAGCA GGTTCGAGC AATGGCGATG ACCACTCAGA AGGAGGCTG 600
 GTGGAGAACC AGTGGACAG CACCATGAAC ATGTTGGGCG GGGAGGCGAG TGCTGGCGG 660
 85 AAGCCCCCTCA AGTCGGGTAT GAAGGAGCTG GCGTGTTC GGGAGAAGGT CACTGAGCAG 720
 CACCGGCGA TGGGCAAGGG TGGCAAGCAT CACTTGGGCC TGGAGGAGCC CAAGAAGCTG 780
 CGACCAACCC CTGCCAGGAC TCCCTGCCAA CAGGAACCTG ACCAGTCTCT GGAGCGGATC 840

WO 02/086443

PCT/US02/12476

5
10
TCCACCATGC GCCTTCCGGA TGAGCGGGGC CCTCTGGAGC ACCTCTACTC CCTGCACATC 900
CCCAACTGTG ACAGCATGAG CCTGTACAAC CTCAAACAGT GCAAGATGTC TCTGAACGGG 960
CAGCGTGGGG AGTGTCTGGTG TGTGAACCCC AACACCCGGA AGCTGATCCA GGGAGCCCCC 1020
ACCATCCGGG GGGACCCCGA GTGTCTATCT TTCTACAATG AGCAGCAGGA GGCTTGGCGG 1080
GTGCACACCC AGCGGATGCA GTAGACCGCA GCCAGCCGGT GCCTGGCGCC CTGCCCCCCC 1140
GCCCTCTCC AACACCCGC AGAAAACGGA GAGTGTCTGG GTGGTGGGTG CTGGAGGATT 1200
TTCCAGTTCT GACACACGTA TTTATATTG GAAAGAGACC AGCACCGAGC TCGGCACCTC 1260
CCCGGCCTCT CTCTCCCGAG CTGCAGATGC CACACCTGCT CCTTCTTGCT TCCCCCGGGG 1320
GAGGAAGGGG GTTGTGGTCG GCGAGCTGGG GTACAGGTTT GGGGAGGGGG AAGAGAAAT 1380
TTTATTTTGG AACCCCTTGG TCCCTTTTGC ATAAGATTAA AGGAAGGAAA AGT

Seq ID NO: 369 Protein sequence
Protein Accession #: NP_000588

15
20
1 11 21 31 41 51
MLPRVGCPL PLPPPLLLPL LPLLLLLLGA SGGGGGARAE VLFRCPPCTP ERLAACGPPP 60
VAPPAVAAV AGGARMPCAE LVREPGCGCC SVCARLEGEA CGVITPROGQ GLRCYPHFGS 120
ELPLQALVMG EGTCEKRRDA EYGASPEQVA DNGDDHSEGG LVENHVDSTM NMLGGGGSAG 180
RKPLKSGMKE LAVFREKVTE QHRQMGKGGK HHLGLEEPKK LRPPPARTPC QQELDQVLER 240
ISTMRLPDER GPLEHLYSLH IPNCDKHGLY NLKQCKMSLN GQRGBCWCVN PNTGKLIQGA 300
PTIRGDPBCH LFYNEQQEAC GVHTQRMQ

25
Seq ID NO: 370 DNA sequence
Nucleic Acid Accession #: NM_004264
Coding sequence: 6-440

30
35
40
1 11 21 31 41 51
GGAACATGCG GGATCGGCTC ACGCAGCTTC AGGACGCTGT GAATTCGCTT GCAGATCAGT 60
TTTGTAAATGC CATTGGAGTA TTGCAGCAAT GTGTCTCTCC TGCTCTTTTC AATAATATTC 120
AGACAGCAAT TAACAAGAC CAGCCAGCTA ACCCTACAGA AGATATGCCC CAGCTTTTTG 180
CAGCACTGAT TGACAGAACA GCAAAAGACA TTGATGTTTT GATAGATTCC TTACCCAGTG 240
AAGAATCTAC AGCTGCTTGA CAGGCTGCTA GCTTGATATA GCTAGAAGAA GAAACCCATG 300
AAGCTGCTAC ATGTGTGGAG GATGTTGTTT ATCGAGGAGA CATGCTTCTG GAGAAGATAC 360
AAAGCCACT TGCTGATATT GCACAGTCAC AGCTGAAGAC AAGAAGTGGT ACCCATAGCC 420
AGTCTCTTCC AGACTCATAG CATCAGTGGG TACCATGTGG CTGAGAAAAG AACTGTTTGA 480
GTGCCATTAA GAATCTGCA TCAGACTTAG ATACAAGCCT TACCAACAAT TACAGAAACA 540
TTAAACACTA TGACACATTA CCTTTTTCAG TATTTTAAAT AGTCTTCTAT TTTCACTCTT 600
GATAAGCTTA TAAATCATGA TTGAATCAGC TTTAAAGCAT CATACCATCA TTTTTTAACT 660
GAGTGAATTT ATTAAGGCAT GTAATACATT AATGAACATA ATATAAGGAA ACATATGTAA 720
AATCTGTGTA TGACATAATT TATGTCTCCA TTTTGTGTGA TTGCCAGTA CTTTACAAT 780
C

45
Seq ID NO: 371 Protein sequence
Protein Accession #: NP_004255

50
1 11 21 31 41 51
MADRLTQLQD AVNSLADQFC NAIGVLQCG PPASFNNIQT AINKDQPNP TREYAQLFAA 60
LIARTAKDID VLIDSLPSEE STAALQAASL YKLEENHEA ATCEDVVYR GDMLEKIQS 120
ALADIAQSQL KTRSGTHSQS LPDS

55
Seq ID NO: 372 DNA sequence
Nucleic Acid Accession #: AJ271091
Coding sequence: 1-1113

60
65
70
75
1 11 21 31 41 51
ATGGAGAATC AGGTGTGAC GCCGCATGTC TACTGGGCTC AGCGACACCG CGAGCTATAT 60
CTGCGGCTGG AGCTGAGTGA CGTACAGAAC CCTGCCATCA GCATCACTGA AAACGTGCTG 120
CATTTCAAG CTCAGGACA TGGTGCCAAA GGAGACAATG TCTATGAATT TCACCTGGAG 180
TTCTTAGACC TTGTGAAACC AGAGCCTGTT TACAACTGA CCCAGAGGCA GGTAAACATT 240
ACAGTACAGA AGAAAGTGAG TCAGTGGTGG GAGAGACTCA CAAAGCAGGA AAAGCGACCA 300
CTGTTTTTGG CTOCTGACTT TGATCGTTGG CTGGATGAAT CTGATGCGGA AATGGAGCTC 360
AGAGCTAAGG AAGAAGAGCG CCTAAATAAA CTCGACTGCG AAAGCGAAGG CTCTCCTGAA 420
ACTCTTACAA ACTTAAGGAA AGGATACCTG TTTATGTATA ATCTTGTCGA ATTCTTGGGA 480
TTCTCCTGGA TCTTTGTCAA CCTGACTGTG CGATTCTGTA TCTTGGGAAA AGAGTCTCTT 540
TATGACACAT TCCTACTGTT GGCAGCATG ATGTATTCTT GCCAGATGCT GGCAGTTGTG 600
GAAACTATCA ATGCAGCAAT TGGAGTCACT ACCTCACCGG TGCTGCCTTC TCTGATCCAG 660
CTTCTTGGAA GAAATTTTAT TTTGTTATC ATCTTTGGCA CCATGGAAGA AATGCGAAGC 720
AAAGCTGTGG TTTTCTTTGT GTTTATTATG TGGAGTGCAA TTGAAATTTT CAGGTACTCT 780
TTCTACATGC TGAAGTGCAAT TGACATGGAT TGGAGGTGTC TCACATGSGT TCGTTACACT 840
CTGTGGATTC CCTTATATCC ACTGGGATGT TTGGCGGAAG CTGTCTCAGT GATTCACTCC 900
ATTCCAATAT TCAATGAGAC CGGACGATTC AGTTTACAT TGCCATATCC AGTGAAATC 960
AAAGTTAGAT TTTCTTTTTC TCTTCAGATT TATCTTATAA TGATATTTT AGGTTTATAC 1020
ATAAATTTTC GTCACTTTTA TAAACAGCGC AGACTGAAAA TGAGGCGCAGG CGCAGTGGCT 1080
CATGCTGTGG ATCCAGCGC TTTGGGAGGC TGA

80
Seq ID NO: 373 Protein sequence
Protein Accession #: CAB69070

85
1 11 21 31 41 51
MENQVLTPHV YWAQRHRELY LRVELSDVQN PAISITENVL HFKAQGHGAK GDNVYEFHLE 60
FLDLVKPEPV YKLTQRQVNI TVQKVSQWV ERLTKQEKRP LFLAPDFDRW LDESDAEMEL 120
RAKEERLINK LRLSEGSPE TLTLNRKGYL FMYNLVQFLG FSWIFVNLTV RFLCILKESF 180

WO 02/086443

PCT/US02/12476

YDTFHTVADM MYFCQMLAVV ETINAAIGVT TSPVLPSLIQ LLGRNFILFI IPGTMEEMQN 240
KAVVFFVFFYL WSAIEIFRYS FYMLTICIDMD WKVLTWLRYT LWIPLYPLGC LAEAVSVIQS 300
IPIFNETGRF SFTLPYPVKI KVRFSFPLQI YLIMIFLGly INFRHLYKQR RLKMRAGAVA 360
HACDPSALGG

Seq ID NO: 374 DNA sequence
Nucleic Acid Accession #: NM_016395
Coding sequence: 1-1113

1 11 21 31 41 51
| | | | |
ATGGAGAATC AGGTGTTGAC GCCGCATGTC TACTGGGCTC AGCGACACCG CGAGCTATAT 60
CTGGCGCTGG AGCTGAGTGA CGTACAGAAC CCTGCCATCA GCATCACTGA AAACGTGCTG 120
CATTTCAAAG CTCAGGACCA TGGTGCCAAA GGAGACAATG TCTATGAATT TCACCTGGAG 180
TTCTTAGACC TTGTGAAACC AGAGCCTGTT TACAACTGA CCCAGAGGCA GGTAACATT 240
ACAGTACAGA AGAAAGTGAG TCAGTGGTGG GAGAGACTCA CAAAGCAGGA AAAGCGACCA 300
CTGTTTTTGG CTCTGACTT TGATCGTTGG CTGAGTGAAT CTGATGCGGA AATGGAGCTC 360
AGAGCTAAGG AAGAAGAGCG CCTAAATAAA CTCCGACTGG AAAGCGAAGG CTCTCTGAA 420
ACTCTTACAA ACTTAAGGAA AGGATACCTG TTTATGTATA ATCTTGGAAT ATTCTTGGGA 480
TTCTCTGGA TCTTTGTCAA CCTGACTGTG CGATTCTGTA TCTTGGGAAA AGAGTCTTTT 540
TATGACACAT TCCATACTGT GGCTGACATG ATGTATTCTT GCCAGATGCT GGCAGTTGTG 600
GAAACTATCA ATGCAGCAAT TGGAGTCACT ACCTCACCAG TGCTGCCTTC TCTGATCCAG 660
CTTCTTGGAA GAAATTTTAT TTTGTTTATC ATCTTTGGCA CCATGGAAGA AATGCAGAAC 720
AAAGCTGTGG TTTCTTTTGT GTTTTATTTG TGGAGTGCAA TTGAAATTTT CAGGTACTCT 780
TTCTACATGC TGACGTGCAT TGACATGGAT TGGAAAGTGC TCACATGGCT TCCTTACACT 840
CTGTGGATTC CCTATATCC ACTGGGATGT TTGGCGGAGG CTGTCTCAGT GATTCACTCC 900
ATTCCAATAT TCAATGAGAC CGGACGATTC AGTTTCACAT TGCCATATCC AGTGAATAATC 960
AAAGTTAGAT TTTCTTTTAT TCTTCAGATT TATCTTATAA TGATATTTT AGGTTTATAC 1020
ATAAATTTTC GTCACTTTTA TAAACAGCGC AGACTGAAAA TGAGGGCAGG CGCAGTGGCT 1080
CATGCTGTG ATCCACAGCG TTTGGGAGGC TGA

Seq ID NO: 375 Protein sequence
Protein Accession #: NP_057479

1 11 21 31 41 51
| | | | |
MENQVLTPHV YWAQRHRELY LRVELSDVQN PAISITENVL HFKAQGHGAK GDNVYEFHLE 60
FLDLVKPEPV YKLTGRQVNI TVQKKVSQWV ERLTKQEKRP LPLAPDFDRW LDESDAEMEL 120
RAKEEERLNK LRLESEGSPE TLNLRKGYL FMYNLVQFLG FSWIFVNLTV RFIILGKESF 180
YDTFHTVADM MYFCQMLAVV ETINAAIGVT TSPVLPSLIQ LLGRNFILFI IPGTMEEMQN 240
KAVVFFVFFYL WSAIEIFRYS FYMLTICIDMD WKVLTWLRYT LWIPLYPLGC LAEAVSVIQS 300
IPIFNETGRF SFTLPYPVKI KVRFSFPLQI YLIMIFLGly INFRHLYKQR RRRYGRKKR 360
STKKDLDFG LPV

Seq ID NO: 376 DNA sequence
Nucleic Acid Accession #: NM_005987
Coding sequence: 1-270

1 11 21 31 41 51
| | | | |
ATGAATTCTC AGCAGCAGAA GCAGCCTTGC ACCCCACCCC CTCAGCCTCA GCAGCAGCAG 60
GTGAAACAAC CTTGCCAGCC TCCACCCAGG GAACCATGCA TCCCCAAAAC CAAGGAGCCC 120
TGCCAAACCA AGGTGCTGTA GCCCTGCCAC CCAAGATGTC CTGAGCCCTG CCAGCCCAAG 180
ATTCCAGAGC CCTGCCAGCC CAAGGTGCCT GAGCCCTGCC CTTCACCGGT CACTCCAGCA 240
CCAGCCAGCA AGAAGACCAA GCAGAAGTAA

Seq ID NO: 377 Protein sequence
Protein Accession #: NP_005978

1 11 21 31 41 51
| | | | |
MNSQQQKQPC TPPPQPPQQQ VRQPQPPPPQ EPCIPKTKEP CQPKVPEPC PKVPSPQPK 60
IPEPCQKVP EPCPSTVTPA PAQKTKQK

Seq ID NO: 378 DNA sequence
Nucleic Acid Accession #: NM_002105
Coding sequence: 74-505

1 11 21 31 41 51
| | | | |
ACAGCAGTTA CACTGCGGCG GCGTCTGTT CTAGTGTGG AGCGTCTGT CTTCACCGGT 60
CTACCTCGCT AGCATGTGCG GCCGCGGCAA GACTGGCGGC AAGGCCGCG CCAAGGCCAA 120
GTGCGGCTCG TCGCGCGCGG GCCTCCAGTT CCCAGTGGGC CGTGACACC GGCTGCTGCG 180
GAAGGCGCCAC TACGCGGAGC GCGTTGGGCG CGGCGCGCCA GTGTACTTGG CGGCACTGCT 240
GGAGTACCTC ACCGCTGAGA TCCCTGGAGT GCGCGGCAAT GCGGCCGCG ACAACAAGAA 300
ACGCGGAATC ATCCCCCGCC ACCCTGCACT GGCCATCCGC AACGACGAGG AGCTCAACAA 360
GCTGCTGGGC GCGGTGACGA TCGCCAGGG AGCGTCTGT CCAACATCC AGGCGTGTCT 420
GCTGCCAAG AAGACCAGCG CCACCGTGGG GCGGAAGCG CCTCTGGGCG GCNAGAAGGC 480
CACCCAGGCT TCCAGAGAG ACTAAGAGGG CCGCGCGCGC GGCGGCGCG CCCAGTCCC 540
CATGCCACCA CAAAGGCCTT TTTAAGGGCC ACCACCGCCC TCATGGAAAG AGCTGAGCCG 600
CTTCAGACTC CCGGCGAAGC GGGCGCGGCG TCCTTCCCT TCCCTCCCT TCGCCCGCT 660
TGGCGCGCGG GCTCTGAGTC CCGCGCGGCG CCGCTCCCG TCCGCGACCG CTTCCGCGAT 720
CGGCTCGGG CCGTCCCTGT CCGCGCTCCG CCTCCGCTA GGGTTCGGGC CTTCCGCGAT 780
CGGCTTGGCG GCTCTCGGG GACCTCCGT GCGCGAAGA CCGGAGCTG CCGGGGGGAG 840

5
10
15
20
25
30
35
40
45
50
55
60
65

```

GCGGCGCGG CCGCACTGCG CCGCTCGGC GTTCGTGACT CAGCCGCCCC ATCCCAGTGC 900
GCTAAGGGGC TGGGGGAGG CCGCAGCACC TTCTGGAAGA CTGGGCTTTC CGCTCTGAGG 960
CAGGCGCGAG GTGGGCGAGT CAGGCGGAGA GCGGCGGCGC CTGAAGGTGA GTGAGGCCCT 1020
CGCGAGCTGC AGCCGCGGTG TCTGTACCC CCGCGGCGTG GTGCTTAGCC CAGGACTTTC 1080
AGAOGCGCGC TGGCGGCGAG GCTTTGGTGG GAGAGACGGG ATGCGCGATT TCGGTCTGGC 1140
GCGCCTTCTG CCGCGGCGAG CCAGGCCTTT CACATCAGCT CTCCTCCAT CTTCATTTCAT 1200
AGGCTCTGCG TGGGCGCGG ACGAAGCACT TGGTAACAGG CACATCTTCC TCCCGAGTGA 1260
CTGCTCTCTA GGAGGACATT TAGGGGAGGG CAGAGGCGTG CAGTTTGGCT TCACGGCTGG 1320
CTATGTGAGC AGCAAGAGTC GTTTGCGGA ACGGACTGG CAGCCAGGCC TGTGGGCCCC 1380
CCGACGCGCG CCCATTTCCC TTCCAGCAAA CTCAACTCGG CAATCCAAGC ACCTAGATAC 1440
CAGCAAGAT CGGTTAATCC CTGTCTGGAC TGAGCCTCCG TTGGCTTCTG AACTGGAATT 1500
CTCAGCTAA CCCTTCCAG ACTAGAACCT TAGGCATTGG GGAGTTTAG ATGGAATAAT 1560
TTTATTAAAG GATTGTTTT TTTT

```

Seq ID NO: 379 Protein sequence
Protein Accession #: NP_002096

1 11 21 31 41 51
20 MSGRGKTGGK ARAKAKSRSS RAGLQFPVGR VHRLLRKGHY AERVGAGAPV YLAALVLEYLT 60
AEILELAGNA ARDNKRTRII PRHLQLAIRN DESLNKLGG VTIAQGGVLP NIQAVLLPKK 120
TSATVGPAPK SGGKKATQAS: QEY

Seq ID NO: 380 DNA sequence
Nucleic Acid Accession #: AL136942
Coding sequence: 184-864

30
35
40
45
50
55
60
65

```

1 11 21 31 41 51
ACGGCTCCGG CAGAAGCTCG GAGCTCTCGG GGTATCGAGG AGGCAGGCCC GCGGGCGCAC 60
GGGCGAGCGG GCGGGAGGCC GGAGCGGCGG AGGAGCGGCG AGCAGCGGCG CCGCGGGCTC 120
CAGGCGAGCG GGTGAGCGCT CCTGAAACT TGGCGCGCGC CTGCGCCAC TCGCGCCGGA 180
GCGATGAAGA TGGTCCGCC CTGAGACGCG TTCTACTCCA ACAGCTGTCT GTTGTGCTGC 240
CATGTCCGCA CCGGCACCAT CCGCTCGCGG GTCTGCTATC TGATCATCAA TGCTGTGGTA 300
CTGTGTGATT TATTGAGTGC CCGGATCAGT ATAACCTTTC AAGTTCTGAA 360
CTGGGAGGTG ACTTTGAGTT CATGGATGAT GCCAACATGT GCATTGCCAT TGGGATTCT 420
CTTCTCATGA TCTGATATG TGCTATGGCT ACTTACGGAG CGTACAAGCA ACGCGCAGCC 480
TGGATCATCC CATCTCTCTG TTACCAAGAT TTTGACTTTG CCTGAACAT GTTGGTTCGA 540
ATCAGCTGCG TTATTATTC AAATCCATT CAGGAATACA TACGGCAACT GCCTCCTAAT 600
TTTCCCTACA GAGATGATGT CATGTCAGTG AATCCTACCT GTTGTGCTCT TATTATCTT 660
CTGTTTATTA GCATTATCTT GACTTTTAAG GGTACTTGA TTAGCTGTGT TTGGAAGTGC 720
TACGATACA TCAATGGTAG GAATCCTCT GATGTCTGCG TTTATGTTAC CAGCAATGAC 780
ACTACGGTGC TGCTACCCCC GTATGATGAT GCCACTGTGA ATGGTGTCTG CAAGGAGCCA 840
CGGCCACCTT ACGTGTCTGC CTAAGCCTTC AAGTGGGCGG AGCTGAGGCG AGCAGCTTGA 900
CTTTGCAGAC ATCTGAGCAA TAGTTCTGTT ATTTCACTTT TGCCATGAGC CTCTCTGAGC 960
TTGTTGTGTT CTGAAATGCT ACTTTTAAAT ATTTAGATGT TAGATTGAAA ACTGTAGTTT 1020
TCAACATATG TTTGCTAGA ACACGTGAT AGATTAACTG TAGAATCTCT CCGTACGAT 1080
TGGGATATA ACGGCTTCA CTAACCTTCC CTAGGCATTG AAACCTTCCC CAAATCTGAT 1140
GGACCTAGAA GTCTGCTTTT GTACCTGCTG GCGCCCAAAG TTGGGCAATT TTCTCTCTGT 1200
TCCCTCTCTT TTGAAATGT AAAATAAAAC CAAAAATAGA CAACTTTTT TCCAGCCATT 1260
CCAGCATAGA GAACAAACCT TTATGGAAC AGGAATGTCA ATTTGTGAAT CATTGTCTTA 1320
ATTAGTAAA TAGAAGTCTT TATGTATGTG TTACAAGAAAT TTCCCCACA ACATCCTTPT 1380
TGACTGAAGT TCAATGACAG TTTGTGTTTG GTGTAAGAGG ATTTCTCCA TGGCCTGAAT 1440
TAAGACCATT AGAAAGCACC AGGCGGTGGG AGCAGTGACC ATCTACTGAC TGTCTTGTG 1500
GATCTTGTGT CCAGGGACAT GGGGTGACAT GCCTCGTATG TGTAGAGGG TGGAAATGGAT 1560
GTGTTTGGCG CTGCATGGGA TCTGGTGCCC CTCTCTCTCT GGATTCACAT CCCCACCCAG 1620
GGCGGCTTTT TACTAAGTGT TCTGCCCTAG ATTGGTTCAA GGAGGTCAAT CCACTGACTT 1680
TATCAAGTGG AATTGGGATA TATTGTATAT ACTTCTGCTT AACCAATGCG AAAAGGGTTT 1740
TCTTTTCCCT GCAAGCTACA TCCTACTGCT TTGAACCTCC AAGTATGTCT AGTCACCTTT 1800
TAAAGTAA ACATTTTCAG AAAAATGAGG ATTGCCCTCC TTGTATGGCG TTTTACCTT 1860
GACTACCTGA ATTGCAAGGG ATTTTATAT ATTCAATGT TACAAAGTCA GCAACTCTCC 1920
TGTGGTTCA TTATTGAATG TGCTGTAAT TAAGTCGTTT GCAATTAATA CAAGGTTTGC 1980
CCACATCCAA AAAAAAAAAA AAAA

```

Seq ID NO: 381 Protein sequence
Protein Accession #: CAB66876

70
75

```

1 11 21 31 41 51
MKWVAPWTRF YSNSCLCCH VRTGTILLGV WYLIINAVVL LILLSALADP DQYNFSSSEL 60
GGDFEFMDDA NMCIATIAISL LMILICAMAT YGAYKQRAAW IIPFFCYQIF DFLNMLVAI 120
TVLIYPSNIQ EYIRQLPPNF PYRDDVMSVN PTCLVLIILL FISIIITFKG YLISCVWNCY 180
RYINGRNSSD VLVVYTSNDT TVLLPPYDDA TVNGAAKEPP PPTVSA

```

Seq ID NO: 382 DNA sequence
Nucleic Acid Accession #: NM_002510
Coding sequence: 92-1774

80
85

```

1 11 21 31 41 51
CAGATGCCAG AAGAACAATG TTGCTCTTGG TGGACGGGCC CAGAGGAATT CAGAGTTAAA 60
CCTTGAGTGC CTGCGTCCGT GAGAATTTCAG CATGGAATGT CTCTACTATT TCCTGGGATT 120
TCTGCTCTCT GCTGCAAGAT TGCCACTTGA TGCGGCCAAA CGATTTCATG ATGTGCTGGG 180
CAATGAAGA CCTTCTGCTT ACATGAGGGA GCACAATCAA TTAATGGCTT GGTCTCTGTA 240
TGAAATGAC TGAATGAAA AACTCTACCC AGTGTGGAAG CGGGGAGACA TGAGGTGGAA 300
AAACTCTGAG AAGGAGGCC GTGTGCAGGC GGTCTGTACC AGTGACTCAC CAGCCCTCTG 360

```

5 GGGCTCAAA ATAACATTG CGGTGAACCT GATATTCCT AGATGCCAAA AGGAAGATGC 420
 CAATGGCAAC ATAGTCTATG AGAAGAAGCTG CAGAAATGAG GCTGGTTTAT CTGCTGATCC 480
 ATATGTTTAC AACTGGACAG CATGGTCAGA GGACAGTGAC GGGGAAAATG GCACCGGCCA 540
 AAGCCATCAT AACGCTCTCC CTGATGGGAA ACCTTTTCTC CACCACCCCG GATGGAGAAG 600
 ATGGAATTTC ATCTACGCTC TCACACACT TGGTCAGTAT TTCCAGAAAT TGGGACGATG 660
 TTCAGTGAGA GTTCTGTGTA ACACAGCCAA TGTGACACTT GGGGCTCAAC TCATGGAAGT 720
 GACTGTCTAC AGAAGACATG GACGGGCATA TGTTCGCATC GCACAAGTGA AAGATGTGTA 780
 CGTGGTAACA GATCAGATTCT CTGTGTTTGT GACTATGTTT CAGAAGAAGC ATCGAAATTC 840
 ATCCGACGAA ACCTTCCTCA AAGATCTCCC CATTATGTTT GATGTCTCTA TTCATGATCC 900
 10 TAGCCACTTC CTCAATTATT CTACCATTAA CTACAAGTGG AGCTTCGGGG ATAATACTGG 960
 CCTGTTTGTG TCCACCAATC ATACTGTGAA TCACACGTAT GTGCTCAATG GAACCTTCAG 1020
 CCTTAACCTC ACTGTGAAGG CTGCAGCACC AGGACCTTGT CCGCCACCCG CACCACCACC 1080
 CAGACCTTCA AACCCACCCC CTCTTTTAGG ACCTGCTGGT GACAAACCCC TGGAGCTGAG 1140
 TAGGATTCCT GATGAAAATC GCCAGATTAA CAGATATGGC CACTTTCAGC CCACCATCAC 1200
 15 AATTGTAGAG GGAATCTTAG AGGTTAACAT CATCCAGATG ACAGACGTCC TGATGCCGGT 1260
 GCCATGGCCT GAAAGCTCCC TAATAGACTT TGTGCTGACC TGCCAAGGGA GCATTCACCAC 1320
 GGAGGTCTGT ACCATCATTT CTGACCCAC CTGCGAGATC ACCCAGAACA CAGTCTGCAG 1380
 CCTGTGATAT TGGGATGAGA TGTGTCTGCT GACTGTGAGA CGAACCTTCA ATGGGTCTGG 1440
 GACGTACTGT GTGAACCTCA CCTTGGGGGA TGACACAAGC CTGGCTCTCA CGAGCACCCT 1500
 20 GATTCTGTGT CTGACACAG ACCCAGCCTC GCCTTAAAGG ATGGCAAACA GTGCCCTGAT 1560
 CTCGTTGGC TCGTTGGCCA TATTGTGAC TGTGATCTCC CTCTTGGTGT AAAAAAACA 1620
 CAAGGAATAC AACCCAATAG AAAATAGTCC TGGGAATGTG GTCAGAAGCA AAGSCCTGAG 1680
 TGTCTTTCTC AACCGTGCAA AAGCCGTGTT CTTCGGGGA AACCAAGAAA AGGATCCGCT 1740
 ACTCAAAAAC CAGAATTTGA AAGGAGTTTC TTAATTTTCG ACCTTGTTC TGAAGCTCAC 1800
 25 TTTTCAGTGC CATTGATGTA AGATGTGCTG GAGTGGCTAT TAACCTTTT TTCTAAAGA 1860
 TTATTGTTAA ATAGATATTG TGGTTTGGGG AAGTTGAATT TTTTATAGT TAAATGTCAT 1920
 TTTAGAGATG GGGAGAGGGA TTATCTGCA GGCAGCTTCA GCCATGTTGT GAAACTGATA 1980
 AAGCAACTT AGCAAGGCTT CTTTTCATTA TTTTATATGT TTCATTATA AAGTCTTAGG 2040
 TAACTAGTAG GATGAAGAAC CTGTGTCCTG AGAGTAAGGA GAGAAGCTAC TATTGATTAG 2100
 30 AGCCTAACCC AGGTTAACTG CAAGAAGAGG CGGGATACIT TCAGCTTTC ATGTAACCTG 2160
 ATGCATAAAG CCAATGTAGT CCAGTTTCTA AGATCATGTT CCAAGCTAAC TGAATCCAC 2220
 TTCAATACAC ACTCATGAAC TCCTGATGGA ACAATAACAG GCCCAAGCCT GTGGTATGAT 2280
 GTGCACACTT GCTGACACTA GAAAAAATAC TACTCTCATA AATGGGTGGG AGTATTTTGG 2340
 TGACAACTTA CTTTCTTGG CTGAGTGAAG GAATGATATT CATATATTTA TTTATTCCAT 2400
 35 GGACATTAG TTAGTGTCTT TTATATACCA GGCATGATGC TGAGTGACAC TCTTGTGTAT 2460
 ATTTCCAAAT TTTGTATAG TCGCTGCACA TATTTGAAT CATATATTAA GACTTTCCAA 2520
 AGATGAGTCT CCTGTTTCTT CATGGCACTG TGATCAGTAA GGATTTCACT TCCTTTTGTG 2580
 ACTAAACCA CTACTATATG GTTAGACATG ACATTCTTTT TCTCTCTTC CTGAAAAATA 2640
 AAGTGTGGGA AGAGACAAA AAAAAAAA

Seq ID NO: 383 Protein sequence
Protein Accession #: NP_002501

45 1 11 21 31 41 51
 MECLYYFLGP LLLAARLPLD AAKRFHDVLG NERPSAYMRE HNQLNGWSSD ENDWNEKLVP 60
 VVKRGMWRWK NSWKGGRVQA VLTSDSPALV GSNITFAVNL IPRCQKEDA NGNIVYEKNC 120
 RNEAGLSADP YYVNWTAWE DSDGENTGQ SHENVFPDGK PFPHPGWRR WNFYVFFHTL 180
 50 GQYFQKLGRG SVRVSVNTAN VTLGPQLMEV TVYRRHGRAY VPIAQVKDVI VVTDQIPFV 240
 TMFQKRNDS SDETLFKDLP IMFVLIHDP SHFLNYSTIN YKNSFGDNTG LFSVSTNHTVN 300
 HTYVLNGTFS LNLTVKAAAP GPCPPPPPPP RPSKPTPSLG PAGDNPLELS RIPDENCOIN 360
 RYGHQATIT IVEGILEVNI IQMTDVLMPV PWPESLIDF VVTCQGSIPT EVCTIISDPT 420
 CEITQNTVCS PVDVDEMCLL TVRRTFNGSG TYCVNLTLGD DTSALSTSL ISVFDKDPAS 480
 55 PLRMANSLI SVGCIAFVT VISLLVYKXH KEYNPINSP GNVVRSKGLS VPLNRKAVF 540
 FPGNQEKDPL LKNQEPKGVG

Seq ID NO: 384 DNA sequence
Nucleic Acid Accession #: NM_001134
Coding sequence: 48-1877

60 1 11 21 31 41 51
 TCCATATTGT GCTTCCACCA CTGCCAATAA CAAAATAACT AGCAACCATG AAGTGGGTGG 60
 AATCAATTTT TTTAATTTTC CTACTAAATT TTAAGTATC CAGAACACTG CATAGAAATG 120
 65 AATATGGAAT AGCTTCCATA TTGGATTCTT ACCAATGTAC TGCAGAGATA AGTTTAGCTG 180
 ACCTGGCTAC CATATTTTTT GCCCAGTTTG TTCAAGAAGC CACTTACAAG GAAGTAAGCA 240
 AAATGGTGAA AGATGCATTG ACTGCAATTG AGAAACCCAC TGGAGATGAA CAGTCTTCAG 300
 GGTGTTTAGA AAACCAAGCTA CCTGCCTTTC TGGAAAGAACT TTGCCATGAG AAAGAAATTT 360
 70 TGGAGAAGTA CGGACATTCA GACTGCTGCA GCCAAAGTGA AGAGGGAAGA CATAACTGTT 420
 TTCTTGCACA CAAAAGGCC ACTCCAGCAT CGATCCCACT TTTCCAAGTT CCAGAACCTG 480
 TCACAAGCTG TGAAGCATAT GAAGAAGACA GGGAGACATT CATGAACAAA TTCAATTATG 540
 AGATAGCAAG AAGGCATCCC TTCTGTATG CACTTACAT TCTTCTTTGG GCTGCTCGCT 600
 ATGACAAAAT AATTCCATCT TGCTGCAAG CTGAAATGTC AGTTGAATGC TTCCAAACAA 660
 75 AGGCAGCAAC AGTTACAAAA GAATTAAGAG AAAGCAGCTT GTTAAATCAA CATGCAATGTG 720
 CAGTAATGAA AATTTTGGG ACCCGAACTT TCCAAGCCAT AACTGTACT AAACAGTCTC 780
 AGAAGTTTAC CAAAGTTAAT TTTACTGAAA TCCAGAAACT AGTCTGGAT GTGGCCCATG 840
 TACATGAGCA CTGTTGCAAG GGAGATGTGC TGGATTGTCT GCAGGATGGG GAAAAATCA 900
 TGTCTCATAT ATGTTCTCAA CAAGACACTC TGTCAACAAA AATAACAGAA TGTGCAAC 960
 80 TGACCAAGCT GGAAGCTGGT CAATGTATAA TTATGTCAGA AATGATGAA AAACCTGAAG 1020
 GTCTATCTCC AATCTAAAC AGGTTTCTAG GAGATAGAGA TTTTAAACCA TTTTCTCAG 1080
 GGGAAAAAAA TATCTTCTTG GCAAGTTTGT TTTATGAATA TTCAAGAAGA CATCTCTCAG 1140
 TTGCTGTCTC AGTAATCTTA AGAGTTGCTA AAGGATACCA GGAGTTATTT GAGAAGTGT 1200
 TCCAGACTGA AATCCCTCTT GAATGCCAAG ATAAAGGAGA AGAAGATTA CAGAAATACA 1260
 85 TCCAGGAGAG CCAAGCATTT GCAAGCGGAA GCTCGGCGCT CTTCAGAAA CATGAGAGAT 1320
 ATTACTTACA AATGCGTTT CTGTTGCTT ACACAAAGAA AGCCCCCAG CTGACCTCGT 1380
 CGGAGCTGAT GGCCATCACC AGAAAAATGG CAGCCACAGC AGCCACTTGT TGCCAACTCA 1440
 GTGAGGACAA ACTATTGGCC TGTGGCGAGG GAGCGGCTGA CATTATTATC GGACACTTAT 1500

GTATCAGACA TGAAATGACT CCAGTAAACC CTGGTGTGG CCAGTGTGC ACTTCTTCAT 1560
 ATGCCAACAG GAGGCATGCT TTCAGCAGCT TGGTGGTGGG TGAACATAT GTCCTCTCTG 1620
 CATCTCTGA TGACCAAGTTC ATTTTCCATA AGGATCTGTG CCAAGCTCAG GGTGTAGCGC 1680
 TGCAACAGAT GAAGCAAGAG TTTCTCATT ACCTTGTGAA GCAAAAGCCA CAAATAACAG 1740
 AGGAACAAC TTAGGCTGTC ATTGCAGATT TCTCAGGCT GTTGGAGAAA TGCTGCCAAG 1800
 GCCAGGAACA GGAAGTCTGC TTGCTGAAG AGGGACAAAA ACTGATTTC AAAAACTCGT 1860
 CTGCTTTGG AGTTTAAAT ACTTCAGGGG AAGAGAAGAC AAAACGAGTC TTTCATTGG 1920
 TGTGAACCTT TCTCTTAAAT TTTAACTGAT TTAACACTTT TTGTGAATTA ATGAAATGAT 1980
 AAAGACTTTT ATGTGAGATT TCTTATCAC AGAAAAAAA TATCTCCAAA TG

Seq ID NO: 385 Protein sequence
 Protein Accession #: NP_001125

1 11 21 31 41 51
 MKWVESIFLI FLNFTBSRT LHRNEYGIAS ILDSYQCTAB ISLADLATIF PAQFVQEATY 60
 KEYSKMKVKA LTAIEKPTGD EQSSGCLENG LPAPLEELCH EKEILEKYGH SDCCSQSEEG 120
 RHNCFLAHRK TPASIPLPQ VPEPVTSCBA YEEDRETFMN KFIYIARRH PFLYAPTILL 180
 WAARYDKIIP SCCKAENAVE CFQTKAATVT KELRESSLLM QHACAVMKNP GTRTFQAITV 240
 TKLSQKFTKV NFEIQLKLV DVAHVHEHCC RGDVLDCLAD GEKIMSYICS QQDILSNKIT 300
 ECKKLTTLER QQCIHAEND EKPEGLSPNL NRFLGDRDFN QFSSGEKNIF LASFVHEYSR 360
 RHPQLAVSVI LRVAQYQEL LEKCFQTEPN LECQDKGEE LQKYIQESQA LAKRSCGLFQ 420
 KLGEYVLQNA FLVAYTKKAP QLTSSSELMAL TRKMAATAAT CCQLSEDKLL ACGEAADI 480
 IGHLCIRHEM TPNVPGVGQC CTSSYANRRP CFSSLVVDET YVPPAFSDDK FIFHKDLQQA 540
 QGVALQTMKQ EFLINLVKQK PQITEEQLEA VIADFSGLLE KCCQGGQEV CFAEGQKLI 600
 SKTRALGV

Seq ID NO: 386 DNA sequence
 Nucleic Acid Accession #: NM_002205.1
 Coding sequence: 1..3149

1 11 21 31 41 51
 ATGGGGAGCC GGACGCGAGA GTCCCCCTCT CACGCCGTGC AGCTGCGCTG GGGCCCCCGG 60
 CGCCGACCCC CGCTSSGTGC GCTGCTGTTG CTGCTSSGTG CGCGCCACC CAGGCTCGGG 120
 GGCTTCAACT TAGACGCGGA GGCOCAGCA GTACTCTCGG GGCOCGCGG CTCTCTCTTC 180
 GGATTCTCAG TGGAGTTTAA CCGCCGCGGA ACAGACGCGG TCAGTGTGCT GGTGGGAGCA 240
 CCCAAGGCTA ATACCAAGCA GCCAGGAGTG CTGCAGGCTG GTGCTGTCTA CCTCTGTCTC 300
 TGGGGTGCCA GCCCACACA GTGCACCCCT ATTGAATTTG ACAGCAAGG CTCTCGGCTC 360
 CTGGAGTCTT CACTGTCCAG CTCAGAGGGA GAGGAGCCTG TGGAGTACAA GTCTCTGCAG 420
 TGGTTCCGGG CAACAGTTTG AGCCCATGGC TCCTCCACTT TGGCATGCGC TCCACTGTAC 480
 AGCTGGCGCA CAGAGAAGGA GCCACTGAGC GACCCCGTGG GCACCTGCTA CCTCTCCACA 540
 GATAACTTCA CCGAATTCTT GGAGTATGCA CCGTCGCGCT CAGATTTCAG CTGGGCGAGA 600
 GGACAGGGTT ACTGCCAAGG AGGCTTCAGT GCCGAGTTCA CCAAGACTGG CCGTGTGGTT 660
 TTAGGTGAGC CAGGAAGCTA TTCTGGCAA GGCAGATCC TGCTGTCCAC TCAGGAGCAG 720
 ATTGCAGAA CTATTACCC CGAGTACCTG ATCAACCTGG TTCAGGGGCA GCTGCAGACT 780
 CGCCAGGCGA GTTCCATCTA TGATGACAGC TACCTAGGAT ACTCTGTGGC TGTGTGTGAA 840
 TTCAGTGTGG ATGACACAGA AGACTTTGTT GCTGTGTGTC CCAAGGGGAA CCTCACTTAC 900
 GGCTATGTCA CCATCTTAA TGGCTCAGAC ATTCGATCCC TCTACAATT CTCAGGGGAA 960
 CAGATGGGCT CTACTTTGG CTATGCAGTG GCCGCCACAG ACCTCAATGG GGACGGGCTG 1020
 GATGACTTGC TGGTGGGGCT ACCCCTGCTC ATGGATGGGA CCGCTGACGG GCGGCTCAG 1080
 GAGGTGGGCA GGGTCTACGT CTACCTGAGC CACCCAGCCG GCATAGAGCC CAGGCCAC 1140
 CTTACCTTCA CTGGCCATGA TGAGTTTGGC GATTTTGGCA GCTCCTGAC CCGCTGGGG 1200
 GACCTGGACC AGGATGCGTA CAATGATGTG GCCATCGGGG CTCCCTTTGG TGGGGAGACC 1260
 CAGCAGGGAG TAGTGTTTGT ATTTCTGGG GGCACAGGAG GGCTGGGCTC TAAGCCTTCC 1320
 CAGGTTTTCG AGCCCTGTG GGCAGCCAGC CACACCCAG ACTTCTTTGG CTCTGCCCTT 1380
 CGAGGAGGCG GAGACCTGGA TGGCAATGGA TATCCTGATC TGATTGTGGG GTCTTTGGT 1440
 GTGGACAAGG CTGTGGTATA CAGGGGCGGC CCGATCGTGT CCGCTAGTGC CTCCCTAC 1500
 ATCTTCCCGC CATGTTCCTA CCGAGAGGAG CGGAGCTGCA GCTTAGAGGG GAACCTGTG 1560
 GCCTGCATCA ACCTTAGCTT CTGCCCTAAT GCTTCTGGAA AACACGTGTC TGACTCCATT 1620
 GGTTCACAG TGGAACCTCA GCTGGAAGTG CAGAAGCAGA AGGGAGGGGT ACGGCGGGCA 1680
 CTGTTCTGG CCTCCAGGCA GGCAACCTG ACCCAGACCC TGCTCATCCA GAATGGGGCT 1740
 CGAGAGGATT GCAGAGAGAT GAAGATCTAC CTCAGGAACG AGTCAGAATT TCAGACAAA 1800
 CTCTCGCGCA TTCATCGC TCTCAACTTC TCCTTGGACC CCAAGCCCC AGTGACAGC 1860
 CAGGCTTCA GGCAGCCCT ACATTATCAG AGCAAGAGCC GGATAGAGGA CAAGGCTCAG 1920
 ATCTTGTGCG ACTGTGGAGA AGACAACATC TGTGTGCTG ACCTGCAGCT GGAAGTGT 1980
 GGGGAGCAGA ACCATGTGTA CCTGGGTGAC AAGAATGCC TGAACCTCAC TTTCCATGCC 2040
 CAGAAATGTG GTGAGGGTGG GGCCTATGAG GCTGAGCTTC GGGTCACCGC CCCTCCAGAG 2100
 GCTGAGTACT CAGGACTCGT CAGACACCCA GGAACCTTCT CCAGCCTGAG CTGTGACTAC 2160
 TTTGCGGTGA ACCAGAGCCG CCGCTGCTG TGTGACCTGG GCAACCCAT GAAGGAGGGA 2220
 GCCAGTCTGT GGGTGGGCT TCGGTTTACA GTCCCTCATC TCGGGGACAC TAAGAAAACC 2280
 ATCCAGTTTG ACTTCCAGAT CCTCAGCAAG AATCTCAACA ACTCGCAAG CGACGTGGTT 2340
 TCCTTTCCGC TCTCCGTGGA GGCTCAGGCC CAGGTCAACC TGAACGCTGT CTCAAGCCT 2400
 GAGGCGAGTG TATTCCAGT AAGCGACTGG CATCCCGAG ACCAGCCTCA GAAGGAGGAG 2460
 GACCTGGGAC CTGCTGTCCA CCATGTCTAT GAGCTCATCA ACCAAGGCCC CAGCTCCATT 2520
 AGCCAGGGTG TGCTGGAAT CAGCTGTCCC CAGGCTCTGG AAGGTCAGCA GCTCTTAT 2580
 GTGACCGAG TTAAGTGAAT CAACCTGACC ACCAATCAC CCATTAAACC AAAGGGCTG 2640
 GAGTTGGATC CCGAGGGTTC CTGACCCAC CAGCAAAAC GGAAGCTCC AAGCCGAGC 2700
 TCTGCTTCT CGGACCTTCA GATCCTGAAA TGCCCGAGG CTGAGTGTTC CAGGCTGCGC 2760
 TGTGAGCTCG GGCCTCTGCA CCAACAAGAG AGCCAAAGTC TGCACTTGCA TTTCCAGTCC 2820
 TGGGCCAAGA CTTTCTTGA GCGGGAGCAC CAGCCATTTA GCCTGCAGTG TGAGGCTGTG 2880
 TACAAGCCC TGAAGATGCC CTACCGAATC CTGCTCGGC AGCTGCCCA AAAAGAGCGT 2940
 CAGGTGGCCA CAGCTGTGCA ATGGAACAG GCAGAAGGCA GCTATGGCT CCCTCTGTG 3000
 ATCATCATCC TAGCCATCCT GTTTGGCCTC CTGCTCCTAG GTCTACTCAT CTACATCCTC 3060
 TACAAGCTTG GATTCTTCAA ACGCTCCCTC CCATATGGCA CCGCATGGA AAAAGCTCAG 3120
 CTCAGGCTC CAGCCACCTC TGATGCTGTA

Seq ID NO: 387 Protein sequence
Protein Accession #: NP_002195.1

```

5      1      11      21      31      41      51
|      |      |      |      |      |
MGSRTPEPPL HAVQLRWGPR RRPPLPLPLL LLLPPPPRVG GFNLDAEAPA VLSGPPGSFF 60
GFSVEFYRPG TDGVSVLVGA PKANTSQPGV LQGGAVYLCF WNASPTQCTP IEFDSKGSRL 120
LESSLSSEEG EEPVEYKSLQ WFGATVRAHG SSILACAPLY SMRTEKEPLS DPVGTCTYLSL 180
10 DNFTRILEVA PCRSDFSWAA GQGYCQGGPS AEFTKTGRVV LGGPGSYFWQ GQILSATQEQ 240
IAESYYPEYL INLVQQLQOT RQASSIYDDS YLGYSVAUGE FSGDDTEDFV AGVPKGNLTY 300
GYVTILNGSD IRSLYNFSGE QMASYFGYAV AATDVNGDGL DDLVVGAPLL MDRTFQGRPQ 360
EVGRVYVYLQ HPAGIEPTPT LTLTGHDEFG RFGSSLTFLG DLDQDGYNDV AIGAPFGGET 420
QQGVVVFVPG GPGGLGSKPS QVLQPLWAAS HTDFFFGSAL RGRDLDGNG YPDLIVGSFG 480
15 VDKAVVYRGR PIVSASASLT IFPAMFNPEE RSCSLEGNFV ACINLSFCLN ASGKHVADSI 540
GFTVELQLDW QKQKGGVRRR LFLASRQATL TQTLIIQNGA REDCREMKIY LRNSEFRDK 600
LSPIHIALNF SLDPQAPVDS HGLRPLALHYQ SKSRIEDKQA ILLDCGEDNI CVPDLQLEVF 660
GEQNHYVYLD KNALNLTFHA QNVGEGGAYE AELRVTAPE AEYSGLVRRH GNFSSLSQDY 720
FAYNQSRLLV CDLGNPMKAG ASLWGLRPT VPHLRDTKKT IQDFQILSK MLNNSQSDVV 780
20 SFRLSVEAQA QVTLNVSQTP EAVLFVPSDW HPRDQPKKEE DLGPAVHVY ELINQGPSSI 840
SQGVLELSCP QALEGQQLLY VTRVTGLNCT TNHPINPKGL ELDPGSLH H QOKREAPSR 900
SASSGPQILK CPBAECFRLR CELGPLHQQE SQSLQLHFRV WAXTFLOREH QPFSLQCEAV 960
YKALMPYRI LPRQLPKQR QVATAVQWTK AEGSYGVPLW IILAILFGL LLLGLLIYIL 1020
YKLGFFKRSI PYGTAMEKAQ LKPPATSDA

```

Seq ID NO: 388 DNA sequence
Nucleic Acid Accession #: NM_002425
Coding sequence: 26..1453

```

30      1      11      21      31      41      51
|      |      |      |      |      |
AAAGAAGGTA AGGGCAGTGA GAATGATGCA TCTTGCAATC CTGTGCTGT TGTGTCTGCC 60
AGTCTGCTCT GCTATTCCTC TGAGTGGGGC AGCAAAAGAG GAGGACTCCA ACAAGGATCT 120
TGCCACAGCA TACCTAGAAA AGTACTACAA CCTCGAAAGG GATGTGAAAC AGTTTAGAAG 180
35 AAAGACAGT AATCTCATTT TTAATAAAT CCAAGGAATG CAGAAGTTCC TTGGGTGAGT 240
GGTGACAGGG AAGCTAGACA CTGACACTCT GGAGGTGATG CGCAAGCCCA GGTGTGGAGT 300
TCTGACGCTT GGTCACTTCA GCTCCTTTCC TGGCATGCCG AAGTGGAGGA AAACCCACCT 360
TACATACAGG ATTGTGAATT ATACACCAGA TTTGCCAAGA GATGCTGTTG ATTCTGCCAT 420
TGAGAAAGCT CTGAAAGTCT GGGAAAGAGT GACTCCACTC ACATTCTCCA GGCTGTATGA 480
40 AGGAGAGGCT GATATAATGA TCTCTTTCG AGTTAAAGAA CATGGAGACT TTTACTCTTT 540
TGATGGCCCA GGACACAGTT TGGCTCATGC CTACCCACCT GGACCTGGGC TTTATGGAGA 600
TATTCACTTT GATGATGATG AAAAATGGAG AGAAGATGCA TCAGGCACCA ATTTATTCCT 660
CGTTGCTGCT CATGAACCTG GCCACTCCCT GGGGCTCTTT CACTCAGCCA ACACCTGAAGC 720
TTTGATGTAC CCACTCTACA ACTCATTAC AGAGCTCGCC CAGTCCGCC TTTCCGAAGA 780
45 TGATGTGAAT GGCATTCACT CTCTCTACGG ACCTCCCTCT GCCTCTACTG AGGAACCCCT 840
GGTGCCCAAA AAATCTGTTT CTTCGGGATC TGAGATGCCA GCCAAGTGTG ATCTCTGCTT 900
GTCTCTCGAT GCCATCAGCA CTCGAGGGGG AGAATATCTG TTTCTTAAAG ACAGATATTT 960
TTGGCGAAGA TCCCACTGGA ACCCTGAACC TGAATTTTCA TTGATTTCTG CATTTTGGCC 1020
CTCTCTTCCA TCATATTTGG ATGCTGCATA TGAAGTTAAC AGCAGGGACA CCGTTTTTAT 1080
50 TTTTAAAGCA AATGAGTTCT GGGCCATCAG AGGAATAGAG GTACAAGCAG GTTATCCAAG 1140
AGGCATCCAT ACCCTGGGTT TTCTCTCAAC CATAAGGAAA ATTGATGCAG CTGTTTCTGA 1200
CAAGGAAAGC AAGAAACAT ACTTCTTTGC AGCGGACAAA TACTGGAGAT TTGATGAAAA 1260
TAGCCAGTCC ATGGAGCAAG GCTTCCCTAG ACTAATAGCT GATGACTTTC CAGGAGTTGA 1320
GCCTAAGGTT GATGCTGAT TACAGGCATT TGGATTTTTC TACTTCTTCA GTGGATCATC 1380
55 ACAGTTTGA TTTGACCCCA ATGCCAGGAT GGTGACACAC ATATTAAAGA GTAACAGCTG 1440
GTTACATTGC TAGGCGAGAT AGGGGGAAGA CAGATATGGG TGTTTTAAAT AAATCTAATA 1500
ATTATTCATC TAATGTATTA TGAGCCAAAA TGGTTAATTT TTCTGTCATG TTCTGTGACT 1560
GAAGAAGATG AGCCTTGCAG ATATCTGCAT GTGTCATGAA GAATGTTTCT GGAATCTTTC 1620
ACTTGCTTTT GAATTCACCT GAACAGAAAT AAGAAATACT CATGTGCAAT AGGTGAGAGA 1680
60 ATGTATTTC ATAGATGTGT TATTACTTCC TCAATAAAAA GTTTTATTTT GGGCCTGTTC 1740
CTT

```

Seq ID NO: 389 Protein sequence
Protein Accession #: NP_002416

```

65      1      11      21      31      41      51
|      |      |      |      |      |
MHLAFVLVLC LPVCSAYPLS GAAKEEDSNK DLAQOYLEKY YNLEKDVQKF RPKDSNLIVK 60
KIQGMQKFLG LEVTGKLDTD TLEVVRKPRC GVPDVGHFSS FPGMPKWRKT HLYTRIVNYT 120
70 PDLPRDAVDS AIEKALKVWE EVTPLTF SRL YEGEADIMIS FAVKEHGDFY SFDGPGHSLA 180
HAYPPGPGLY GDHIFDDDEK WTEDASGTNL FLVAAHELGH SLGLFHSANT EALMPLYLNS 240
PTELAQFRLS QDDVNGIQSL YGPPFPASTE PLVPTKSVPS GSEMPAKCDP ALSFDAISTL 300
RGEYLFFKDR YFWRSHWNP EPEFHLISAF WPSLPSYLD AYEVSNDTV FIFKGNFPA 360
75 IRGNEVQAGY PRGIHTLGF PTIRKIDAAV SDKEKKKTYF FAADKYWRF ENSQSMQGF 420
PRLIADDPFG VEPKVDVQLQ AFGFFYFFSG SSQFEFDPNA RMVTHILKSN SWLHC

```

Seq ID NO: 390 DNA sequence
Nucleic Acid Accession #: NM_002421.2
Coding sequence: 1..1409

```

80      1      11      21      31      41      51
|      |      |      |      |      |
ATGCACAGCT TTCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTG ACACAGCTTC 60
CCAGCGACTC TAGAAACACA AGAGCAGAT GTGGACTTAG TCAGAAATA CCTGGAAAAA 120
85 TACTACACC TGAAGAATGA TGGGAGGCAA GTTGAAGAAG GGAGAAATAG TGGCCAGTGG 180
GTTGAAAAAT TGAAGCAAT GCAGGAATTC TTTGGGCTGA AAGTACTGG GAAACAGAT 240
GCTGAAACCC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCTGATGT GGCTCAGTTT 300

```

5
10
15
20

GTCTCTACTG AGGGGAACCC TCGCTGGGAG CAAACACATC TGACCTACAG GATTGAAAAT 360
TACACGCCAG ATTITGCCAAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTTCCTCAACT 420
TGGAGTAATG TCACACCTCT GACATTCAAC AAGGTCTCTG AGGGTCAAGC AGACATCATG 480
ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCTCT TTGATGGACC TGGAGGAAAT 540
CTTGCTCATG CTTTTCAACC AGGCCAGGT ATTGGAGGGG ATGCTCATT TTGATGAAGAT 600
GAAAGGTGGA CCAACAATT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGAACTC 660
GGCCATTCTC TTGGACTCTC CCATTCTACT GATATCGGG CTITGATGTA CCCTAGCTAC 720
ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATT ATGGCATCCA AGCCATATAT 780
GGACGTTCOC AAAATCTGT CCAGCCCATC GGCCCAAAA CCCCAAAAGC ATGTGACAGT 840
AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA 900
TTCTACATGC GCACAAATCC CTTCTACCCG GAAGTTGAGC TCAATTTTCAT TTCTGTTTTC 960
TGGCCCAAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC 1020
CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC 1080
CCCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAAGT TGAAGCATAT CGATGCTGCT 1140
CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTTG CTAACAAATA CTGGAGGTAT 1200
GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTTCT 1260
GGAATGGCC ACAAAGTTGA TGCAGTTTTC ATGAAAGATG GATTTTCTA TTTCTTTTCAT 1320
GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380
AATAGCTGGT TCAACTGCAG GAAAAATTAG.

Seq ID NO: 391 Protein sequence
Protein Accession #: NP_002412.1

25
30
35

1 11 21 31 41 51
MHSFPPLLLL LFWGVVSHSF PATLETQEOD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60
VEKLKQMQEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRWE QTHLTYRIEN 120
YTPDLPRADV DHAIEKAFQL WSNVTPLTFT KVSEGOADIM ISFVRGDHRD NSPFDGPGGN 180
LAHAFQPGPG IGGDAHFDEED ERWNNFREY NLHRVAHAEL GHSGLSHST DIGALMYPST 240
TFSGDVQLAQ DDIDGIAIY GRSONPVQPI GPQTPKACDS KLTFDAITTI RGEVMFFKDR 300
FYMRTNFFYP EVELNFIISV WPQLPNGLEA AYEFADRDEY RFFKGNKYWA VQGVNVLHGY 360
PKDIYSSPGF PRTVKHIDAA LSEENTGKTY FFFVANKYWR YDEYKRSMDPG YPKMIAHDFP 420
GIGHKVDVAF MKDGFYFFH GTRQYKDFK TKRILTLQKA NSWFNCRKN

Seq ID NO: 392 DNA sequence
Nucleic Acid Accession #: NM_002421.2
Coding sequence: 1..1409

40
45
50
55
60
65

1 11 21 31 41 51
ATGCACAGCT TTCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTC ACACAGCTTC 60
CCAGCGACTC TAGAAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAAA 120
TACTACAACC TGAAGAATGA TGGGAGGCAA GTTGAAAAAG CGAGAAATAG TGGCCCAAGT 180
GTTGAAAAAT TGAAGCAAT GCAGGAATTC TTTGGGCTGA AAGTGACTGG GAAACACAGT 240
GCTGAAACCC TGAAGGTGAT GAAGCAGCCC AGATGTGAG TGCCTGATGT GGCTCAGTTT 300
GTCTCTACTG AGGGGAACCC TCGCTGGGAG CAAACACATC TGACCTACAG GATTGAAAAT 360
TACACGCCAG ATTITGCCAAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTTCCTCAACT 420
TGGAGTAATG TCACACCTCT GACATTCAAC AAGGTCTCTG AGGGTCAAGC AGACATCATG 480
ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCTCT TTGATGGACC TGGAGGAAAT 540
CTTGCTCATG CTTTTCAACC AGGCCAGGT ATTGGAGGGG ATGCTCATT TTGATGAAGAT 600
GAAAGGTGGA CCAACAATT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGCCCTC 660
GGCCATTCTC TTGGACTCTC CCATTCTACT GATATCGGG CTITGATGTA CCCTAGCTAC 720
ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATT ATGGCATCCA AGCCATATAT 780
GGACGTTCOC AAAATCTGT CCAGCCCATC GGCCCAAAA CCCCAAAAGC ATGTGACAGT 840
AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA 900
TTCTACATGC GCACAAATCC CTTCTACCCG GAAGTTGAGC TCAATTTTCAT TTCTGTTTTC 960
TGGCCCAAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC 1020
CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC 1080
CCCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAAGT TGAAGCATAT CGATGCTGCT 1140
CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTTG CTAACAAATA CTGGAGGTAT 1200
GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTTCT 1260
GGAATGGCC ACAAAGTTGA TGCAGTTTTC ATGAAAGATG GATTTTCTA TTTCTTTTCAT 1320
GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380
AATAGCTGGT TCAACTGCAG GAAAAATTAG

Seq ID NO: 393 Protein sequence
Protein Accession #: NP_002412.1

70
75
80

1 11 21 31 41 51
MHSFPPLLLL LFWGVVSHSF PATLETQEOD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60
VEKLKQMQEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRWE QTHLTYRIEN 120
YTPDLPRADV DHAIEKAFQL WSNVTPLTFT KVSEGOADIM ISFVRGDHRD NSPFDGPGGN 180
LAHAFQPGPG IGGDAHFDEED ERWNNFREY NLHRVAHAEL GHSGLSHST DIGALMYPST 240
TFSGDVQLAQ DDIDGIAIY GRSONPVQPI GPQTPKACDS KLTFDAITTI RGEVMFFKDR 300
FYMRTNFFYP EVELNFIISV WPQLPNGLEA AYEFADRDEY RFFKGNKYWA VQGVNVLHGY 360
PKDIYSSPGF PRTVKHIDAA LSEENTGKTY FFFVANKYWR YDEYKRSMDPG YPKMIAHDFP 420
GIGHKVDVAF MKDGFYFFH GTRQYKDFK TKRILTLQKA NSWFNCRKN

Seq ID NO: 394 DNA sequence
Nucleic Acid Accession #: NM_014331.2
Coding sequence: 1..1506

85

1 11 21 31 41 51

5
10
15
20
25
30
35
40
45
50
55

ATGGTCAGAA AGCCTGTTGT GTCCACCATC TCCAAAGGAG GTTACCTGCA GGGAAATGTT 60
AACGGGAGGC TGCCCTCCCT GGGCAACAAG GAGCCACCTG GGCAGGAGAA AGTCAGCTG 120
AAGAGGAAG TCACCTTACT GAGGGGAGTC TCCATTATCA TTGGCACCAT CATTGGAGCA 180
GGAAATCTCA TCTCTCTCAA GGGCGTGCTC CAGAACACGG GCAGCGTGGG CATGCTCTCT 240
ACCATCTGGA CGGTGTGTGG GGTCTGTGCA CTAATTGGAG CTTTGTCTTA TGCTGAATTG 300
GGAACAACATA TAAAGAAATC TGGAGGTCAT TACACATATA TTTTGGAAAT CTTTGGTCCA 360
TTACCAAGCTT TTGTACGAGT CTGGGTGGAA CTCCTCATAA TACGCCCTGC AGCTACTGCT 420
GTGATATCC TGACATTGG ACGCTACATT CTGGAAACAT TTTTATTCA ATGTGAATC 480
CCTGAACCTG CGATCAAGCT CATTACAGCT GTGGGCATAA CTGTAGTGAT GGTCTCTAAAT 540
AGCATGAGTG TCAGCTGGAG GSCCGGATC CAGATTTTCT TAACCTTTTG CAAGCTCACA 600
GCAATCTGTA TAATTATAGT CCCTGGAGTT ATGCAGCTAA TTAAGGTCA AACGCAGAAC 660
TTTAAAGAGC GTTTTCAGG AAGAGATTCA AGTATTACGC GGTGGCCACT GGCCTTTTAT 720
TATGGAATGT ATGCATATGC TGGCTGGTTT TACCTCAACT TTGTTACTGA AGAAGTAGAA 780
AACCCGAAAA AAACCATTC CCTTGCATAA TGTATATCCA TGGCCATTGT CACCATTGGC 840
TATGTGCTGA CAATGTGGC TACTTTAAG ACCATTAAAT CTGAGGAGCT GCTGCTTTCA 900
AATGCAGTGG CAGTGACCTT TCTGAGCGG CTACTGGGAA ATTCTCATT AGCAGTTCOG 960
ATCTTTGTTG CCCTCTCCCTG CTTTGGCTCC ATGAACGGTG GTGTGTTTGC TGCTCCAGG 1020
TTATTCTATG TTGCGTCTGG AGAGGGTCAC CTTCCAGAAA TCCTCTCCAT GATTCTATGC 1080
CGCAAGCACA CTCCTCTACG AGCTGTTATT GTTTTGACCC CTTTGACAAT GATAATGCTC 1140
TTCTCTGGAG ACCTCGACAT TCTTTTGAAT TTCTCTAGTT TTGCGAGGTG GCTTTTATT 1200
GGGCTGGCAG TTGCTGGGCT GATTATCTT CGATACAAAT GCCCAGATAT GCATCGTCT 1260
TTCAAGGTGC CACTGTTCAT CCCAGCTTTG TTTTCTTCA CATGCCCTT CATGGTGGCC 1320
CTTTCCCTCT TGGCATGACC ATTTAGTACA GGGATTGGCT TCGTCATCAC TCTGACTGGA 1380
GTCCCTGCGT ATTATCTCTT TATTATATGG GACAAGAAAC CCAGGTGGTT TAGAATAATG 1440
TCAGAGAAAA TAACCAAGAA ATTACAAATA ATACTGGAAG TTGTACCAGA AGAAGATAAG 1500
TTATGAACATA ATGCACTTGA GATCTTGGCA ATCTGCCCAA GGGGAGACAC AAAATAGGGA 1560
TTTTTACTTC ATTTTCTGAA AGTCTAGAGA ATTACAACT TGGTGATAAA CAAAAGGAGT 1620
CAGTTATTTT TATTGATATA TTTAGCATA TTCGAACTAA TTTCTAAGAA ATTTAGTTAT 1680
AATCTATATG AGTTATAGAA AGTGAATATG CAGTTATCTT ATGAGTCGCA CAATCTCTGA 1740
GTCTCTGATA CCTACCTATT GGGGTTAGGA GAAAGACTA GACAATTACT ATGTGGTCAT 1800
TCTCTACAA ATATGTTAGC ACGGCAAAGA ACCTTCAAT TGAAGACTGA GATTTTCTG 1860
TATATATGGG TTTTGTAAAG ATGTTTTCAC ACACACAGA TGTCTATACT GTGAAAAGTG 1920
TTTTCAATTC TGAAAAAAG CATACTCAT GATTATGGCA AAGAGGAGAG AAAGAAATTT 1980
ATTTTACATT GACATTGCTT TGCTTCCCTC TAGATACCAA TTAGATAAC AAACACTCAT 2040
GCITTAATGG ATTATACCCA GAGCACTTTG AACAAAGTTC AGTGGGGATT GTTGAATACA 2100
TTAAAGAAGA GTTCTAGGGG GCTACTGTTT ATGAGACACA TCCAGGAGTT ATGTTTAAAT 2160
AAAAATCCCT GAGAATTTAT TATGTCAGAT GTTTTTCAT TCATTATCAG GAAGTTTATG 2220
TTATCTGTC TTTTCTTTT TACATCAGT TTGATCAGGA AAGTGATATA CACATCTTAG 2280
AGCAAGAGTT AGTTTGGTAT TAAATCTCTA TTAGAACAA CACCTGTTTC ACTAATAACT 2340
TACCCCTGAT GAGTCTATCT AAACATATGC ATTTTAAAGC TTCAAATTAC ATTATCAACA 2400
TGAGAGAAAT AACCAACAAA GAAGATGTTT AAAATAATAG TCCCATATCT GTAATCATAT 2460
CTACATGCAA TGTTAGTAAT TCTGAAGTTT TTTAAATTA TGGCTATTTT TACACGATGA 2520
TGAATTTTGA CAGTTTGTGC ATTTTCTTTA TACATTTTAT ATTCTCTCTG TAAATATCT 2580
CTTCAGATGA AACTGTCCAG ATTAATTAGG AAAAGGCATA TATTAACATA AAAATTGCAA 2640
ATCTAGGCTT TGTCAGTAAT TTCCACACCT TAATTATCAT TCAACTTGCA AAAGAGACAA 2700
CTGATAAGAA GAAAAATTGA ATGAGAATCT GTGGATAAGT GTTTGTGTTG AGAAGATGTT 2760
GTTTGGCCAG TATTAGAAAA TACTGTGAGC CGGGCATGGT GCCTTACATC TGTAATCCCA 2820
GCACCTTGGG AGGCTGAGGG GGTGGATCAC CTGAGGTGGG GAGTCTAGA CCAGCCTGAC 2880
CAACATGGAG AAACCCCATC TCTACTAAAA ATACAAAATT AGCTGGGCAT GGTGGCACAT 2940
GCTGGTAATC TCAGCTATTG AGGAGGCTGA GGCAGGAGAA TTGCTTGAAC CCGGGAGGCG 3000
GAGGTTGTCAG TGAGCCAAAG TTAGCACCAT GTACTCCAGC CTGGGTGACA AAGTCAGACT 3060
CCATCTCCAA AAAAAAATAA AAAA

Seq ID NO: 395 Protein sequence
Protein Accession #: NP_055146.1

60
65
70

1 11 21 31 41 51
MVRKPVVSTI SKGGYLGQNV NGRLPVSLGNK EPPGQEKVQL KRKVTLLRGV SIIIGTIIGA 60
GIFISPKGVL QNTGVSVMGL TIWTVCGVLS LFGALSVAEL GTTIKKSQGH YTYILEVFGP 120
LPAFVRVWVE LLIRPAATA VISLAFGRYT LEPPFIQCEI PELAIKLITA VGITVVMVLN 180
SMSVWSARI QILFTPKLT AILIIIVPGV MQLIKGQTON FKDAFSGRDS SITRLPLAFY 240
YGMVYAGWF YLNFVTEVEE NPEKTIPLAI CISMATIGV VVLNVAYFT TINAEELLS 300
NAVAITFSER LLGNFSLAVP IFVALSCFGS MNGGVFAVSR LFYVASREGH LPEILSMIHV 360
RKHTPLPAVI VLHPLTMI ML PSLDLSLIL FLSFARWLP GLAVAGLIYL RYKCPDMERP 420
FKVPLFIPAL FSFTCLFMVA LSLYSDPFST GIGFVITLTG VPAYYLFIIW DKKPRWFRIM 480
SEKITRTLQI ILEVVPEDK L

Seq ID NO: 396 DNA sequence
Nucleic Acid Accession #: NM_006528
Coding sequence: 57..764

75
80
85

1 11 21 31 41 51
GCCGCCAGCG GCTTTCTCGG ACGCCTTGCC CAGGGGGCGG CCCGACCCCG TGCACCATGG 60
ACCCGCTCG CCCCTGGGG CTGTGATTCT TGCTGCTTTT CCTGACGGAG GCTGCACCTG 120
GCGATGCTGC TCAGAGAGCA ACAGGAAATA ACGGGAGAT CTGTCTCTCT CCCCTAGACT 180
ACGGACCTCG CCGGGCCCTA CTTCTCGGTT ACTACTACGA CAGGTACAGC CAGAGCTGCC 240
GCCAGTTCCT GTACGGGGGC TGCAGGGGCA ACGCCAACAA TTTCTACACC TGGGAGGCTT 300
GCGACGATGC TTGCTGGAGG ATAGAAAAAG TTCCCAAAGT TTGCGGCTG CAAGTGAGTG 360
TGGAACGACA GTGTGAGGGG TCCACAGAAA AGTATTTCTT TAATCTAAGT TCCATGACAT 420
GTGAAAAATT CTTTTCGGGT GGTGTGACCC GGAACCGGAT TGAGAACAGG TTTCCAGATG 480
AAGCTACTTG TATGGGCTTC TGCACACCAA AGAAAAATCC ATCATTTTGC TACAGTCCAA 540
AAGATGAGGG ACTGTCTCT GCGCAATGTGA CTCGCTATTA TTTTAATCCA AGATACAGAA 600
CCTGTGATGC TTTCACTAT ACTGGCTGTG GAGGGAATGA CAATAACTTT GTTAGCAGGG 660

WO 02/086443

PCT/US02/12476

AGGATTGCAA ACGTGCATGT GCAAAAGCTT TGA AAAAGAA AAGAAGATG CCAAAGCTTC 720
 GCTTTGGCCG TAGAATCCGG AAAATTCGGA AGAAGCAATT TTAACATTTC TTAATATGTC 780
 ATCTTGTTTG TCTTTATGGC TTATTTGGCT TTATGGTGTG ATCTGAAGAA TAATATGACA 840
 GCATGAGGAA ACAAATCATT GGTGATTIAT TCACCAAGTTT TTATTAATAC AAGTCACTTT 900
 TCAAAAATT TGGATTTTTT TATATATAAC TAGCTGCTAT TCAATGTGA GTCTACCATT 960
 TTTAATTAT GGTTCAACTG TTTGTGAGAC GAATCTTGGC AATGCATAAG ATATAAAGC 1020
 AAATATGCT CACTCATTTT TTGGGGTGTG ATTCTGTATT TCAGAAGAGG ATCATAACTG 1080
 AAACAACATA AGACAATATA ATCATGTGCT TTTAACATAT TTGAGAATAA AAAGGACTAG 1140
 CC

Seq ID NO: 397 Protein sequence
 Protein Accession #: NP_006519

1 11 21 31 41 51
 MDPARPLGLS ILLFLTEAA LGDAAQEPTG NNABICLLPL DYGPICRALLL RYYDYRYTQS 60
 CRQFLYGGCE GNANFYTWE ACDDACWRIE KVPKVCRLQV SVDDQCEGST EKYFFNLSSM 120
 TCEKFFSGGC HNRNIENRFP DEATCMGFCA PKKIPSPCYG PKDEGLCSAN VTRYFFNPYR 180
 RTCDAPTYTG CGGNDNNEVS REDCKRACAK ALKKKKKKMPK LRFASIRIKI RKKQF

Seq ID NO: 398 DNA sequence
 Nucleic Acid Accession #: NM_001508.1
 Coding sequence: 1..1361

1 11 21 31 41 51
 ATGGGCTTAC CCAGCCTCCC GGGCAGTGAC TGCTCCCAA TCATTGATCA CAGTCATGTC 60
 CCCGAGTTTG AGGTGGCCAC CTGGATCAAA ATCACCTTAA TTCTGGTGA CCTGATCATC 120
 TTCGTGATGG GCCTTCTGGG GAACAGCGTC ACCATTGGG TCACCCAGGT GCTGCAGAAG 180
 AAAGGATACT TGCAGAAGGA GGTGACAGAC CACATGGTGA GTTTGGCTTG CTCCGACATC 240
 TTGGTGTTC TCATCGGCAT GCCCATGGAG TTCTACAGCA TCATCTGGAA TCCCTTGACC 300
 ACGTCCAGCT ACACCCCTGT CTGCAAGCTG CACACTTTCC TCTTCGAGGC CTGCAGCTAC 360
 GCTACGCTGC TGACGCTGCT GACGCTCAGC TTTGAGCGCT ACATCGCCAT CTGTCACCCC 420
 TTCAAGTACA AGGCTGTGTC GGGACCTTGC CAGGTGAAGC TGCTGATTGG CTTCGTCTGG 480
 GTCACCTCCG CCTCGTGGC ACTGCCCCTG CTGTTTGCCA TGGGTACTGA GTACCCCTTG 540
 GTGAACGTGC CCAGCCACCG GGGTCTCACT TGCAACCGCT CCAGCACCGG CCACCAACGAG 600
 CAGCCCGAGA CCTCCAATAT GTCCATCTGT ACCAACCTCT CCAGCCGCTG GACCGTGTTC 660
 CAGTCCAGCA TCTTCGCGCG CTTCGTGCTC TACCTCGTGG TCTGCTCTC CGTAGCCTTC 720
 ATGTGCTGGA ACATGATGCA GGTGCTCATG AAAAGCCAGA AGGGCTCGCT GGC CGGGGGG 780
 ACCGCGCCTC CGCAGCTGAG GAAGTCCGAG AGCGAAGAGA GCAGGACCGC CAGGAGGCAG 840
 ACCATCATCT TCCTGAGGCT GATTGTTGTG ACATTGGCCG TATGCTGGAT GCCCAACCAG 900
 ATTGGAGAGA TCATGGCTGC GGCCTCAACC AAGCAGGACT GGACGAGGTC CTACTTCCGG 960
 GCGTACATGA TCTCTCTCCC CTCTCGGAG ACGTTTTTCT ACCTCAGCTC GGTCACTCAAC 1020
 CCGCTCTCTG ACAGCGTGTG CTGCGAGCAG TTTGCGCGGG TGTTCGTGCA GGTGCTGTGC 1080
 TGCGCGCTGT CGCTGACAGA CGCCAACCAC GAGAAGCGCC TGCGGTGACA TGCGCACTCC 1140
 ACCACGAGCA GCGCCGCTCT TGTGCGCGC CCGTTGCTCT TCGCGTCCCG GCGCCAGTCC 1200
 TCTGCAAGCA GAAGTGTGAA GATTTTCTTA AGCACTTTTC AGAGCGAGGC CGAGCCCCAG 1260
 TCTAAGTCCC AGTCATTGAG TCTCGAGTCA CTAGAGCCCA ACTCAGGCGC GAAACCGACC 1320
 AATCTGCTG CAGAGAATGG TTTTCAGGAG CATGAAGTTT GA

Seq ID NO: 399 Protein sequence
 Protein Accession #: NP_001499.1

1 11 21 31 41 51
 MASPSLPGSD CSQIIDHSHV PEFEVATWIK ITLILVYLII FVMGLLGNV TIRVTQVLQK 60
 KGYLQKEVD HMVSLACSDI LVFLIGMPME FYSIIWNPLT TSSYTLCKL HTFLFEACSY 120
 ATLLHLVLTLS FERYIAIChP FRYKAVSGPC QVKLLIGFVW VTSALVALPL LFAMGTEYPL 180
 VNVPSEHRLT CNRSSTRHHE QPETSNSMIC TNLSSRWTFV QSSIFGAFVW YLVVLLSVAF 240
 MCWMMQVLM KSQKSLAGG TRPPQLRKSE SEESRTARRQ TIIFLRLIVV TLAVCWMPNQ 300
 IRRIMAAAKP KHDWTRSYFR AYMILLPFSE TFFYLSSVIN PLLYTVSSQQ FRRVFVQVLC 360
 CRLSLQHANH EKRLRVHAHS TTDSARFVQR PLLFASRRQS SARRTEKIFL STFQSEAEQ 420
 SKSQSLLES LEPNSGAKPA NSAAENGFBQ HEV

Seq ID NO: 400 DNA sequence
 Nucleic Acid Accession #: NM_006475.1
 Coding sequence: 28..2538

1 11 21 31 41 51
 AACAGAACTG CAACGGAGAG ACTCAAGATG ATTCCTCTTT TACCCATGTT TTCTCTACTA 60
 TTGCTGCTTA TTGTTAACCC TATAAACGCC AACAAATCATT ATGACAAGAT CTGCTGCTAT 120
 AGTCGTATCA GGGGTGCGGA CCAAGGCCCA AATGCTGTG CCCTTCAACA GATTTTGGGC 180
 ACCAAAAAGA AATACTTCAG CACTTGTAAG AACTGGTATA AAAAGTCCAT CTGTGGACAG 240
 AAAACGACTG TTTTATATGA ATGTTGCCCT GGTATATGA GAATGGAAGG AATGAAAGGC 300
 TGCCCGAGCAG TTTTGCCCAT TGACCATGTT TATGGCACTC TGGGCATCGT GGGAGCCACC 360
 ACAACGCAGC GCTATTCTGA GCCTCAAAA CTGAGGGAGG AGATCGAGGG AAAGGGATCC 420
 TTAATTAAT TTTGACCGAG TAATGAGGCT TGGGACAAC TGGATTCTGA TATCGTAGA 480
 GGTTTGGAGA GCAACGTGAA GTTTGAATTA CTGAATGCTT TACATAGTCA CATGATTAAT 540
 AAGAGAAATG TGACCAAGGA CTAAAAAAT GGCATGATTA TTCCTCAAT GTATAACAAT 600
 TTGGGGCTTT TCATTAAACCA TTATCTTAAT GGGGTGTGCA CTGTTAATTG TGCTCGAATC 660
 ATCCATGGGA ACCAGATTGC AACAAATGGT GTTGTCCATG TCATTGACCG TGTGCTTACA 720
 CAAATTTGTA CCTCAATTCA AGACTTCATT GAAGCAGAAG ATGACCTTTC ATCTTTTAGA 780
 GCAGCTGCCA CCAATCGGGA CATATTGGAG GCCCTTGGAA GAGACGGTCA CTTCACTC 840
 TTTGCTCCCA CCAATGAGGC TTTTGAGAAA CTTCCAGCAG GTGTCTTAGA AAGGTTCTAG 900
 GGAGACAAAG TGGCTTCCGA AGCTCTTATG AAGTACCACA TCTTAATATC TCTCCAGTGT 960
 TCTGAGTCTA TTATGGGAGG AGCAGTCTTT GAGACGCTGG AAGGAAATAC AATTGAGATA 1020

5
10
15
20
25
30
35
40
Seq ID NO: 401 Protein sequence
Protein Accession #: NP_006466.1

1 11 21 31 41 51
| | | | |
MIPFLPMFSL LLLLIVNPIN ANNHDKILA HSRIRGRDQG PNVCALQQIL GTKKKYFSTC 60
KNWYKKSICG QKTTVLVECC PGYMRMEGMK GCPAVLPIDH VYGLTGIVGA TTTQRYSDAS 120
KLREEIEGKG SFTYFAPSNE AWDNLDSDIR RGLSNVNVVE LLNALHSHMI NKRLMTKDLK 180
NGMIIPSMYN NLGLFPIHYV NGVVTVNCAR I IHGNQIATN GVHVHIDRVL TOIGTSIQDF 240
IEAEDDLSSF RAAAITSDIL EALGRDGHFT LPAPTNEAFE KLPGRVLERF MGDKVASEAL 300
MKYHILNTLQ CSSEIMGGAV FETLEGNTIE IGCDDGDSITV NGIKMVNKKD IVTNNGVIHL 360
IDQVLIPOSA KQVIELAGKQ QTTFTDLVAQ LGLASALRPD GEYTLAPVN NAFSDDTLSM 420
VQRLKLLILO NHLLKVKVGL NELYNGQILE TIGGKQLRVF VYRTAVCIEN SCMEKSKQGG 480
RNGAHIHIRE IIPKAEKSLH EKLKQDKRFS TFLSLLEAAD LKELLTPQPD WTLFVPTNDA 540
FKGMTSEKEE ILIRDKNALQ NIILYHLTPG VFIGKGFEPG VTNILKTTQG SKIFLKEVND 600
TLVLNELKSK ESIDIMTNGV IHVVDKLLYP ADTPVGNQDL LEILNKLIKY IQIKFVRGND 660
FKRIPVTYVT TKIITKVVEP KIKVIEGSLQ PIKITEGPTL TKVKEGEPE FRLIKEGETI 720
TEVIHGEPII KKYTKIIDGV PVEITEKETR EERIITGPEI KYTRISTGGG ETEETLKLIL 780
QEEVTKVKTF IEQGDGHLFE DEEIKRLLQG DTPVRKLQAN KKVQGSRRRL REGRSQ

60
Seq ID NO: 402 DNA sequence
Nucleic Acid Accession #: NM_002416
Coding sequence: 40..417

1 11 21 31 41 51
| | | | |
ATCCAATACA GGAGTGACTT GGAACCTCAT TCTATCACTA TGAAGAAAAG TGGTGTCTTT 60
TTCTCTCTGG GCATCATCTT GCTGGTTCCT ATTGGAGTGC AAGGAACCCC AGTAGTGAGA 120
AAGGGTCTCG GTTCTCTGAT CAGCACCAAC CAAGGGACTA TCCACCTACA ACTCTTGAAA 180
GACCTTAAAC AATTGGCCCC AAGCCCTTCC TGCAGAGAAA TTGAAATCAT TGCTACACTG 240
AAGAATGGAG TTCAACATG TCTAAACCCA GATTGAGCAG ATGTGAAGGA ACTGATTAAA 300
AAGTGGGAGA AACAGGTCAG CCAAAAGAAA AAGCAAAAGA ATGGGAAAAA ACATCAAAAA 360
AAGAAGTTCG TGAAGTTTCG AAAATCTCAA CGTCTCGTC AAAAGAAGAC TACATAAGAG 420
ACCACTTCAC CAATAAGTAT TCTGTGTTAA AAATGTTCTA TTTTAATTAT ACGCTATCA 480
TTCCAAAGGA GGATGGCATA TAATACAAAG GCTTATTAAT TTGACTAGAA AATTTAAAC 540
ATTACTCTGA AATTGTAAGT AAAGTTAGAA AGTTGATTTT AAGAATCCAA ACGTTAAGAA 600
TTGTTAAAGG CTATGATTGT CTTGTTCCTT CTACACCCCA CCAGTTGAAT TTCAATCATG 660
TTAAGGCCAT GATTTTAGCA ATACCCATGT CTACACAGAT GTTCACCCAA CCACATCCCA 720
CTCACACACG CTGCTCGGAA GAGCAGCCCT AGGCTTCCAC GTACTGCAGC CTCAGAGAG 780
TATCTGAGGC ACATGTCAGC AAGTCTTAAG CCGTGTAGCA TGCTGGTGAG CCAAGCAGTT 840
TGAAATTTAG CTGACCTTCA CCAAGCTGCT GTGGCCATCA ACCTCTGTAT TTGAATCAGC 900
CTACAGGCTC CACACACAAT GTGTCTGAGA GATTCAATGCT GATTGTTATT GGGTATCACC 960
ACTGGAGATC ACCAGTGTGT GGCTTTTACA GCCTCTCTTC TGGCTTTGGA AGCCATGTGA 1020
TTCCATCTTG CCGCTCTCAG CTGACCACTT TATTTCTTTT TGTTCCCTTT TGCTTCATTC 1080
AAGTCAGCTC TTCTCCATCC TACCACAATG CAGTGCCCTT CTCTCTCCCA GTGCACCTGT 1140
CATATGCTCT GATTATCTGT AGTCAACTCC TTCTCTCATCT TGTCGCCAAC ACCCCACAGA 1200
AGTGCTTTCT TCTCCCAAT CATCTCTACT CAGTCCAGCT TAGTTCAGT CCTGCTCTCT 1260
AAATAAACCT TTTTGGACAC ACAAAATTAT TAAAACCTCC TGTTCACCTT GGTTCAGTAC 1320
CACATGGGTG AACACTCAAT GGTTAACATA TTCTTGGGTG TTTATCCTAT CTCTCCAACC 1380

5 AGATTGTCTG CTTCTTGAGG GCAAGAGCCA CAGTATATTT CCTGTCTTCT TCCACAGTGC 1440
 CTAATAATAC TGTGGAACATA GGTTTTAATA ATTTTAAAT TGATGTTGTT ATGGGCAGGA 1500
 TGGCAACCAG ACCATTGTCT CAGAGCAGGT GCTGGCTCTT TCTGGCTAC TCCATGTTGG 1560
 CTAGGCTCTG GTAACCTCTT ACTTATTATC TTCAGGACAC TCACACAGG GACCAGGGAT 1620
 GATGCAACAT CCTGTCTTCT TTATGACAGG ATGTTTGCTC AGCTTCTCCA ACAATAAGAA 1680
 GCACGTGGTA AAACACTTGC GGATATTCTG GACTGTTTTT AAAAAATATA CAGTTTACCG 1740
 AAAATCATAT AATCTTACAA TGAAAAGGAC TTTATAGATC AGCCAGTGAC CAACCTTTTC 1800
 CCAACCATAC AAAAATTCCT TTTCCGAGG GAAAAGGGCT TTCTCAATAA GCCTCAGCTT 1860
 TCTAAGATCT AACAGATAG CCACCGAGAT CCTTATCGAA ACTCATTTTA GGCAATATG 1920
 10 AGTTTTATTG TCCGTTTACT TGTTTCAGAG TTTGTATTGT GATTATCAAT TACCACACCA 1980
 TCTCCCATAG AGAAAGGGAA CGGTGAAGTA CTAAGCGCTA GAGGAAGCAG CCAAGTCGGT 2040
 TAGTGAAGC ATGATTGGTG CCCAGTTAGC CTCTGCAGGA TGTGGAAACC TCCTTCCAGG 2100
 GGAGGTTTCT TGAATTGTGT AGGAGAGGTT GTCTGTGGCC AGAATTTAAA CCTATCTCA 2160
 CTTTCCCAA TTGAATCACT GCTCACACTG CTGATGATT AGAGTGTCTG CCGGTGGAGA 2220
 15 TCCACCCGGA ACGTCTTATC TAATCATGAA ACTCCCTAGT TCCTTCATGT AACTTCCCTG 2280
 AAAAATCTAA GTGTTTCATA AATTGAGAG TCTGTGACCC ACTTACCTTG CATCTCACAG 2340
 GTAGACAGTA TATACTAAC AACCAAGAC TACATATTGT CACTGACACA CACGTTATAA 2400
 TCATTATACA TATATATACA TACATGCATA CACTCTCAA GCAATAAAT TTTCACTTCA 2460
 20 AAACAGTATT GACTTGATA CCTTGTAAAT TGAAATATTT TCTTTGTAA AATAGAATGG 2520
 TATCAATAAA TAGACCATTA ATCAG

Seq ID NO: 403 Protein sequence
Protein Accession #: NP_002407

25 1 11 21 31 41 51
 MKKSGVLFLL GIILLVLIGV QGTEVVRKGR CSCISTNQGT IHLQSLKDLK QFAPSPSCEK 60
 ISIIATLKNG VQTCLNPDSA DVKELIKKWE KQVSQKKKQK NGKHQKHKV LKVRKRSR 120
 QKKTT

Seq ID NO: 404 DNA sequence
Nucleic Acid Accession #: NM_006670
Coding sequence: 85..1347

35 1 11 21 31 41 51
 CCGGCTCGCG CCCTCCGGGC CCAGCCTCCC GAGCCTTCGG AGCGGGCGCC GTCCAGCCCC 60
 AGCTCCGGGG AAACGCGAGC CGCGATGCCT GGGGGGTGCT CCGGGGGCCC GCGCCCGGG 120
 40 GACGGGGCTG TGGGGCTGGC GCGACTAGCG CTGGTACTCC TGGGCTGGGT CTCTCTCTCT 180
 TCTCCCACTT CCTGGGCATC CTCTTCTTCC TCTCTGGGCG GGTCTCTGGC TCCGCGCTG 240
 TCGGCCAGC CCGCGCTGCC GGACCAAGTG CCGCGCTGT GCGAGTGCTC CGAGGCGAGC 300
 CGCACAGTCA AGTGGCTTAA CGCAATCTG ACCGAGGTGC CCACGAGCTC GCGCGCTAC 360
 GTGCGCAACC TCTTCTTAC CGGCAACCAG CTGGCGGTGC TCCCTGCCGG GCGCTTGGCC 420
 CGCGGGCGCG CGCTGGCGGA GCTGGCGCGC CTCAACCTCA GCGGCGCGCG CCTGAGCAG 480
 45 GTGCGCGCGG GCGCCTTCGA GCATCTGCCC AGCTCGACCT AGCTCGACCT CAGCCACAAC 540
 CCACTGGCGG ACCTCAGTCC CTTCGCTTTC TCGGGCAGCA ATGCCAGCT CTGGGCCCC 600
 AGTCCCTTGT TGAACCTGAT CCGAACCAC ATCGTGCCCC CTGAAGATGA GCGGCGAAG 660
 CGGAGCTTGG AGGGCATGTT GGTGGCGGCC CTGCTGGCGG CCGGTGCACT GCAGGGGCTC 720
 CGCGCTTGG AGCTGGCCAG CAACCACTTC CTTTACCTGC CGCGGATGT GCTGGCCCAA 780
 50 CTGCCCAGCC TCAGGCACCT GGACTTAAGT AATAATTGCG TGGTGAGCCT GACCTACGTG 840
 TCCCTCCGCA ACCTGACACA TCTAGAAAGC CTCCACCTGG AGGACAATGC CCTCAAGGTC 900
 CTTCACAATG GCACCTCTGG TGAATTGCAA GGTCTACCCC ACATTAGGTT TTTCTGGAC 960
 AACCAATCCT GGGTCTGCGA CTGCCACATG GCAGACATGG TGACCTGGCT CAAGGAAACA 1020
 55 GAGGTAGTGC AGGCGAAAGA CCGGCTCACC TGTGCATATC CGGAAAAAT GAGGAATCGG 1080
 GTCTCTTGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTCC CCCATCCCTG 1140
 CAAACCTCTT ATGCTCTTCT GGGTATTGTT TTAGCCCTGA TAGGCGCTAT TTTCTCCTG 1200
 GTTTTGTATT TGAACCGCAA GGGGATAAAA AAGTGGATGC ATAACATCAG AGATGCCCTG 1260
 AGGGATCACA TGAAGGGTA TCATTACAGA TATGAAATCA ATCGGAGCCC CAGATTAACA 1320
 AACCTCAGTT CTAACCTGGA TGTCTGAGAA ATATTAGAGG ACAGACCAAG GACAACTCTG 1380
 60 CATGAGATGT AGACTTAAGC TTTATCCCTA CTAGGCTTGC TCCACTTTCA TCCCTCACTA 1440
 TAGATACAAAC GAACTTTGAC TAAAGCAGT GAAGGGGATT TGCTTCTTGT TTATGTAAG 1500
 TTTCTCGGTG TGTCTGTATA ATGTAAGACG ATGAACAGTT GTGTATAGTG TTTTACCTTC 1560
 TTCTTTTCTT TGAACCTCT CAACAGTAT GGAGGGATT TTCAGTTTC AGCATGAACA 1620
 TGGGCTTCTT GCTGTCTGTC TCTCTCTCAG TACAGTTCAA GGTGTAGCAA GTGTACCCAC 1680
 65 ACAGATAGCA TTCAACAAAA GCTGCCTCAA CTTTTTCGAG AAAAACTCTT TATTCATAAA 1740
 TATCAGTTTT ATCTCATGT ACCTAAGTTG TGGAGAAAAT AATTGCATCC TATAAACTGC 1800
 CTGCAGAGT TAGCAGGCTC TTCAAAATAA CTCCATGGTG CACAGGAGCA CCTGCATCCA 1860
 AGAGCATGCT TACATTTTAC TGTCTGTCAT ATTACAAAAA ATAACTTGCA ACTTCATAAC 1920
 70 TTCTTTGACA AAGTAAATTA CTTTTTTGAT TGCAGTTTAT ATGAAATGT ACTGATTTT 1980
 TTTTAATAAA CTGCATCGAG ATCCAACCGA CTGAATTGTT AAAAAAATAA AAAAAATAAG 2040
 ATTCTTAAAA GAA

Seq ID NO: 405 Protein sequence
Protein Accession #: NP_006661

75 1 11 21 31 41 51
 MPGGCSRGPA AGDGLRLAR LALVLLGWVS SSSPTSSASS FSSAPFLAS AVSAQPLPFD 60
 QCPALCESE AARTVKCVNR NLTEVPTDLP AYVRNLFLTG NQLAVLPAGA FARRPPLAEL 120
 80 AALNLSGSR LDEVRAFAFH LPSLRQLDLS HNPDLSPF AFGSNASVS APSPLVELIL 180
 NHIVPPEDER QNRSPEGVVV AALLAGRALQ GLRLLELASN HFLYLPDVL AQPLSLRLHD 240
 LSNLSVSLT YVSRNLTHL ESLHLEDNAL KVLNGTLAE LQGLPHIRVF LDNNPWCDC 300
 EMADMVTLK ETEVQGGKDR LTCAYPEKMR NRVLLELNSA DLDCLPILP SLQTSYVPLG 360
 85 IVALIGAIIF LVLVLYNRKG IKKWMHNRD ACRDHMEGYH YRYEINADPR LTNLSNSNDV

Seq ID NO: 406 DNA sequence
Nucleic Acid Accession #: Eos sequence

Coding sequence: 1..927

1 11 21 31 41 51
5 ATGCTCTGGG GGTGCTCCCG GGGCCCCGCC GCGGGGGAAG GGGGTCTGCG GCTGGGCGCA 60
CTAGCGCTGG TACTCTCTGG CTGGGTCTCC TGCTCTCTCC CCACCTCTCC GGCATCTCTC 120
TTCTCTCTCT CGGCGCGCTT CCTGGCTTCC GCGGTGTCCG CCCAGCCCCC GCTGCCCGAC 180
CAGTGCCCGG CGCTGTGCGA GTGCTCCGAG GCAGCGGCGA CAGTCAAGTG CGTTAACCGC 240
AATCTGACCG AGGTGCCAC GCACCTGCCC GCCTACGTGC GCAACCTCTT CCTTACCGGC 300
10 AACCTGCTGG CCAGCAACCA CTCTCTTTAC CTGCCGCGGG ATGTGCTGGC CCAACTGCCC 360
AGCCTCAGGC ACCTGGACTT AAGTAATAAT TCGCTGGTGA GCCTGACCTA CGTGTCTCTC 420
CGCAACCTGA CACATCTAGA AAGCCTCCAC CTGGAGGACA ATGCCCTCAA GGTCTCTCAC 480
AATGGCAACC TGGCTGAGT GCAAGGTCTA CCCCACATTA GGGTTTTCCT GGACAACAAT 540
15 CCCTGGGTCT GCGACTGCCA CATGGCAGAC ATGGTGACCT GGCTCAAGGA AACAGAGGTA 600
GTGCAGGGCA AAGACCGGCT CACCTGTGCA TATCCGGAAG AAATGAGGAA TCGGGTCTCT 660
TTGGAATCA CAGATGCTGA CCTGGACTGT GACCCGATTG TCCCCCATC CCTGCAAAAC 720
TCTTATGCTT TCCTGGGTAT TGTTTTAGCC CTGATAGGCG CTATTTTCCT CTGGTTTGTG 780
TATTGAACC GCAAGGGGAT AAAAAAGTGG ATGCATAACA TCAGAGATGC CTGCAGGGAT 840
20 CACATGGAAG GGTATCATTA CAGATATGAA ATCAATGCGG ACCCCAGATT AACAAACCTC 900
AGTTCTAACT CGGATGTCTT CGAGTGA

Seq ID NO: 407 Protein sequence
Protein Accession #: Eos sequence

25 1 11 21 31 41 51
MPGGCSRGA AGDGRRLRL LALVLLGWVS SSSPTSSASS FSSAPFLAS AVSAQPLPLD 60
QCPALCESE AARTVVKCVNR NLTEVPTDLP AYVRNLFITG NQLASNHFLY LPRDVLALQL 120
30 SLRHLDSNN SLVSLTYVSF RNLTHLESLH LEDNALKVLH NGTLAEIQLG PHIRVFLDNN 180
PWVCDCHMAD MVTWLKETEV VQKDRILTCA YPEKMRNRVL LEIINSADLDC DPILPPLSLQ 240
SYVFLGIVLA LIGAIPLVL YLNRKGIKKM MENIRDACRD HMEGYHYRYE INADPRLTNL 300
SSNSDVLE

Seq ID NO: 408 DNA sequence
Nucleic Acid Accession #: NM_000095.1
Coding sequence: 26..2299

1 11 21 31 41 51
40 CAGCACCAG CTCCCCGCCA CCGCCATGGT CCCCAGACAC GCTTGGCTTC TTCTGCTCAC 60
CCTGGCTGCC CTGGGCGCT CCGGACAGGG CCAGAGCCCG TTGGGCTCAG ACCTGGGCCC 120
GCAGATGCTT CGGGAATCGC AGGAAACCAA CCGCGCGCTG CAGGACGTGC GGGACTGGCT 180
CGCGCAGCAG GTCCAGGAGA TCACGTTCTT GAAAAACAG GTGATGGAAT GTGACGCGTG 240
CGGGATGCGC CAGTCAGTAC GCACCGGCTT ACCCAGCGTG CCGCCCTGCG TCCACTGCGC 300
45 GCCCGGCTTC TGCTTCCCGG GCGTGGGCTG CATCCAGACG GAGAGCGGCG GCCGCTGCGG 360
CCCCGCCCCC GCGGGCTTCA CCGGCAACCG CTGCGACTGC ACCGACGTCA ACGAGTGCAA 420
CGCCCAACCC TGCTTCCCGG GAGTCCGCTG TATCAACACC AGCCCGGGGT TCCGCTGCGA 480
GGCTTGCCTG CCGGGGTACA CCGGCCCCAC CCACACGGGC GTGGGGCTGG CTTTGGCCAA 540
50 GGCACCAACG CAGGTTTGCA CGACATCAA CGAGTGTGAG ACCGGGCAAC ATAAGTGGCT 600
CCCCAACTCC GTGTGATCA ACACCCGGGG CTCTTCCAG TGCGGCCCTG GCCAGCCCGG 660
CTTCGTGGGC GACGAGCGGT CCGGCTGCCA GCGCGCGCGA CAGCGCTTCT GCCCCGACGG 720
CTGCGCCAGC GAGTGCCAGC AGCATGCAGA CTGCGTCTTA GAGCGCGATG GCTCGGGGCT 780
GTGCGTGTGT CGGTTGGCTT GGGCCGGCAA CCGGATCTCT TGTGTTGCGC ACGACTGACT 840
55 AGACGGCTTC CCGGACGAGA AGCTGCGCTG CCGGAGCGCG CAGTGCCTGA AGGCAAACTG 900
CGTACTGTGT CCCAACTCAG GGCAGGAGGA TGTGGACCGC GATGGCATCG GAGACGCTGT 960
CGATCCGGAT CGCGACGGGG ACGGGGTCCC CAATGAAAAG GACAACTGCC CGCTGCTGCG 1020
GAACCCAGAC CAGCGCAACA CCGACGAGGA CAAGTGGGCG GATGCGTGCG ACAACTGCGC 1080
GTCCCAAGAC AACGACGAGC AAAAGGACAC AGACCAAGGAC GCGCGGGGCG ATGCGTGCGA 1140
60 CGACGACATC GACGCGGACC GGATCCGCAA CCGAGCGGAC AACTGCCCTA GGGTACCCAA 1200
CTCAGACAGG AAGGACAGTG ATGGCGATGG TATAGGGGAT GCCTGTGACA ACTGTCCCAA 1260
GAAGAGCAAC CCGGATCAGG CGGATGTGGA CCACGACTTT GTGGGAGATG CTTGTGACAG 1320
CGATCAAGAC CAGGATGAGG ACGGACATCA GGAATCTCGG GACAACTGTC CCACGCTGCC 1380
TAACAGTGCC CAGGAGGACT CAGACCAAGA TGGCCAGGGT GATGCTGCGG ACGACGACGA 1440
65 CGACAATGAC GGAATCCCTG ACAGTCCGGA CAACTGCCGC CTGTTGCCCTA ACCCCGCGCA 1500
GGAGGACCGG GACAGGAGCG GCGTGGGCGA CGTGTGCCAG GACGACTTTG ATGCAGACAA 1560
GGTGGTAGAC AAGATCGAGC TGTGTCCGGA GAACGCTGAA GTCACGCTCA CCGACTTCAG 1620
GGCCTCCAG ACAGTCTGTC TGGACCCGGA GGGTGACGCG CAGATTGACC CCAACTGGGT 1680
GGTGTCAAC CAGGGAAGGG AGATCTGTGA GACAATGAAC AGCGACCCAG GCGTGGCTGT 1740
70 GGGTTACACT GCCTTCAATG CGGTGGAATT CGAGGGCAGC TTCCATGTGA ACACGCTCAC 1800
GGATGACGAC TATGCGGGCT TCATCTTTGG CTACAGGAC AGCTCCAGCT TCTACGTGGT 1860
CATGTGGAAG CAGATGGAGC AAACGTATTG GCAGGCGAAC CCCTTCCGTG CTGTGGCCGA 1920
GCCTGGCATC CAATCAAGG CTGTGAAGTC TTCCACAGGC CCGGGGGAAC AGCTGCGGAA 1980
CGCTCTGTGG CATAAGGAG ACACAGAGTC CCAGGTGCGG CTGCTGTGGA AGGACCGCGG 2040
75 AAACGTGGGT TGAAGGACA AGAAGTCTTA TCGTTGGTTC CTGCAGCACC GGGCCCAAGT 2100
GGGCTACATC AGGGTGGGAT TCTATGAGGG CCCTGAGCTG GTGGCCGACA GCAAGTGGT 2160
CTTGGACACA ACCATGGGGG GTGGCCGCTT GGGGCTCTTC TGCTTCTCCC AGGAGAACAT 2220
CATCTGGGCC AACCTGGGTT ACCGCTGCAA TGACACCATC CCAGAGGACT ATGAGACCCA 2280
TCAGCTGCGG CAAGCCTAGG GACCAGGGTG AGGACCCGCG GATGACGAGC CACCTCACCC 2340
80 GCGGCTGGAT GGGGGTCTTG CACCCAGCCC AAGGGGTGGC CGTCTGAGG GGGAGTGGAG 2400
AAGGGCTCAG AGAGGACAAA ATAAAGTGTG TGTGAGGG

Seq ID NO: 409 Protein sequence
Protein Accession #: NP_000086.1

85 1 11 21 31 41 51
MVPDTACVLL LTLAALGASG QQSPLGSDL GPQMLRELQE TNAALQDVLD WLRQQVREIT 60

WO 02/086443

PCT/US02/12476

FLKNTVMECD ACQMQQSVRT GLPSVRPLLH CAPGFCPPGV ACIQTESGGR CGPCPAGFTG 120
 NGSHCTDVNE CNAHPCFPRV RCINTSPGFR CEACPPGYSG PTHQGVGLAF AKANKQVCTD 180
 INECETGQIN CVPNSVCINT RGSFQCGPCQ PGFVGDQASG CQRGAQRFCP DGSFSECHEH 240
 ADCVLERDGS RSCVCRVWGA NGILCGRDT DLDGFPDEKL RCPEPQCRKD NCVTVPNSGQ 300
 EDVDRDGDG ACDDPDADGG VPNEKDNCPV VRNPQDRNTD EDKNGDACDN CRSQKNDQDK 360
 DTQDQGRGDA CDDIDIGDRI RNQADNCPRV PMSQKQSDG DGIGDACDNC PQKSNPDQAD 420
 VDHDYVGDAC DSDQDQDGDG HQDSRDNCPT VNSAQEDSD HDGGGDACDD DDDNDGVFDS 480
 RDNCRLVFNF QGEDADRQGV GDVQDDFDA DKVVDKIDVC PENAEVLTLD FRAFQTVVLD 540
 PEGDAQIDPN WVVLNQGREI VQTMNSDPGL AVGYTAENGV DFEFTFHVNT VTDDYAGFI 600
 PGYQSSSFY VVMKQMEQT YWQANPFRV ABPGIQLKAV KSSTGPGEQL RNALWBTGDT 660
 ESSVRLWKD FRNVGKDKK SYRNFLQHRP QVGYIRVRFY EGPELVADSN VVLDITMRGG 720
 RLGVFCFSQE NIWNLRYR CNDTIPEDYE THQLRQA

Seq ID NO: 410 DNA sequence
 Nucleic Acid Accession #: NM_001565.1
 Coding sequence: 67..363

1 11 21 31 41 51
 GAGACATTCC TCAATTGCTT AGACATATTC TGAGCCTACA GCAGAGGAAC CTCAGTCTC 60
 AGCACCATGA ATCAAACTGC GATTCTGATT TGCTGCCTTA TCTTCTGAC TCTAAGTGGC 120
 ATTCAGAGAG TACCTCTCTC TAGAACCGTA CGCTGTACCT GCATCAGCAT TAGTAATCAA 180
 CCTGTTAATC CAAGGTCTTT AGAAAACTT GAAATTATTC CTGCAAGCCA ATTTTGTCCA 240
 CGTGTGAGA TCATTGCTAC AATGAAAAAG AAGGGTGAGA AGAGATGCTT GAATCCAGAA 300
 TCGAAGGCCA TCAAGAAATT ACTGAAAGCA GTTAGCAAGG AAATGTCTAA AAGATCTCCT 360
 TAAAGCCGGA GGGGAGCAAA ATCGATGCAG TGCTTCCAAG GATGGACCAC ACAGAGGCTG 420
 CCTCTCCATC CACTCTCCTA CATGGAGTAT ATGTCAAGCC ATAATTGTTC TTAGTTTGCA 480
 GTTACACTAA AAGGTACCA ATGATGGTCA CCAATCAGC TGCTACTACT CCTGTAGGAA 540
 GGTAAATGTT CATCATCCTA AGCTATTGAG TAATAACTCT ACCCTGGCAC TATAATGTAA 600
 GCTCTACTGA GGTGCTATGT TCTTAGTGGA TGTTCTGACC CTGCTTCAAA TATTTCCCTC 660
 ACCTTTCCCA TCTTCCAAGG GTACTAAGGA ATCTTTCTGC TTGGGGTITT ATCAGAATTC 720
 TCAGAACTC AAATAACTCA AAGGTATGCA ATCAATCTG CTTTTTAAAG AATGCTCTTT 780
 ACTTCATGGA CTTCCTCTGC CATCCTCCCA AGGGGCCCAA ATCTTTTCTG TGGCTACCTA 840
 CATACAATTC CAAACACATA CAGGAAGGTA GAAATATCTG AAAATGTATG TGTAGTATT 900
 CTTATTTAAT GAAAGACTGT ACAAAGTATA AGTCTTAGAT GTATATATTT CCTATATTGT 960
 TTTCACTGTA CATGGAATAA CATGTAATTA AGTACTATGT ATCAATGAGT AACAGGAAAA 1020
 TTTTAAAAAT ACAGATAGAT ATATGCTCTG CATGTTACAT AAGATAAATG TGCTGAATGG 1080
 TTTTCAATA AAAATGAGGT ACTCTCCTGG AAATATTAAG

Seq ID NO: 411 Protein sequence
 Protein Accession #: NP_001556.1

1 11 21 31 41 51
 MNQTALICC LIPLTLGSIQ GVPLSRTVRC TCISISNQPV NPSLEKLEI IPASQFCPRV 60
 EIIATMKKKG EKRCCLNPESK AIKNLLKAVS KEMSKRSP

Seq ID NO: 412 DNA sequence
 Nucleic Acid Accession #: XM_057014
 Coding sequence: 143..874

1 11 21 31 41 51
 GGGAGGGAGA GAGGCGCGCG GGTGAAAGGC GCATTGATGC AGCCTGCGGC GGCTCGGAG 60
 CGCGGCGGAG CCAGAGCGCTG ACCACGTTCC TCTCCTCGGT CTCTCCGCC TCCAGCTCCG 120
 CGCTGCCCCG CAGCGCGGAG CCATGCGACC CCAGGGCCCC GCGCGCTCCC CGCAGGGGCT 180
 CGCGGCGCTC CTGCTGCTCC TGCTGCTGCA GCTGCGCGCG CGCTCGAGCG CCTCTGAGAT 240
 CCCCAGGGGG AAGCAAAAGG CGCAGCTCCG GCAGAGGGAG GTGGTGGACC TGTATAATGG 300
 AATGTGCTTA CAAGGGCCAG CAGGAGTGCC TGGTCGAGAC GGGAGCCCTG GGGCCATGG 360
 CATTCGGGTG ACACCTGGGA TCCCAGGTCT GGATGGATTG AAAGGAGAAA AGGGGGAATG 420
 TCTGAGGAAA AGCTTTGAGG AGTCCTGGAC ACCCAACTAC AAGCAGTGT CATGGAGTTC 480
 ATTGAATTAT GGCATAGATC TTGGGAAAAT TGCGGAGTGT ACATTTACAA AGATGCGTTC 540
 AAATAGTGCT CTAAGAGTGT TGTTCACTGG CTCACCTCG CTAAATGCA GAAATGCATG 600
 CTGTCAAGCT TGGTATTTC CATTCAATGG AGCTGAATGT TCAGGACCTC TTCCATTGA 660
 AGCTATAATT TATTGGACC AAGGAAGCCC TGAAATGAAT TCAACAATTA ATATTCATCG 720
 CACTTCTTCT GTGGAAGGAC TTTGTGAAGG AATTGGTGCT GGATTAGTGG ATGTGTCTAT 780
 CTGGGTGGC ACTTGTTCAG ATTACCCAAA AGGAGATGCT TCTACTGGAT GGAATTCAGT 840
 TTCTCGCATC ATTATTGAAG AACTACCAA ATAAATGCTT TAATTTTCAT TTGCTACCTC 900
 TTTTATTATT ATGCTTGGGA ATGGTTCAC TAAATGACAT TTTAAATAAG TTTATGTATA 960
 CATCTGAATG AAAAGCAAAG CTAATATGT TTACAGACCA AAGTGTGATT TCACACTGTT 1020
 TTTAAATCTA GCATTATTCA TTTTGCTTCA ATCAAAAGTG GTTTCATAT TTTTITTAGT 1080
 TGGTATGAAT ACTTCTCTCA TAGTCACATT CTCTCAACCT ATAATTGGA ATATTGTTGT 1140
 GGTCTTTTGT TTTTCTCTCT AGTATAGCAT TTTTAAAAAA ATATAAAAGC TACCAATCTT 1200
 TGTACAAATT GTAATGTTA AGAATTTTTT TTATATCTGT TAAATAAAAA TTATTTCCAA 1260
 CAACCTTAAA AAAAAAAAAA AAAA

Seq ID NO: 413 Protein sequence
 Protein Accession #: XP_057014

1 11 21 31 41 51
 MRPGPAASP QRLRGLLLLL LLQLPAPSSA SEIPKQKQKA QLRQREVVDL YNGMCLQGPA 60
 GVPGRDGPSP ANGIPGTGPI PGRDGPKEK GECLRESFER SWTPNYKQCS WSSLNYGIDL 120
 GKIAECTFTK MRSNSALRVL FSGSLRLKCR NACQQRWYFT PNGAECSSPL PIEAIYLDQ 180
 GSPENNSTIN IHRSTSSVEGL CEGIGAGLVD VAIWVGTCSD YPKGDASTGW NSVSRIIEE 240
 LPK

WO 02/086443
Seq ID NO: 414 DNA sequence
Nucleic Acid Accession #: XM_084007
Coding sequence: 138..2405

PCT/US02/12476

```
5 1 11 21 31 41 51
CTCGTGGCGA ATTGGGCAAG AGACCGCGTG TTCGCGCGTG GTAGAGATT CTGCAAGACA 60
CCAGTGGGCC CGTGTGGAA CAAACCTGCG CGCGTGGCGG GCGGTGGGA CAAAGAGGCC 120
GCGGAGACGA AGGCGCAATG GCGAGGAAGT TATCTGTAAT CTTGATCCTG ACCCTTGGCC 180
10 TCTCTGTAC AAATCCCTT CATGAAC TAAAGCAGTCG TTTCCCCAG ACCACTGAGA 240
AAATTAGTCC GAATGGGAA TCTGGCATT ATGTGACTT GGCAATTTC ACACGGCAAT 300
ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAAATAA TTCTTTGTCA GTTGAAGGGT 360
TCAGAAAATT ACTTCAAAT ATAGGCATAG ATAAGATTAA AAGAAATCAT ATACACCATG 420
ACACGACCA TCACCTCAGC CACGAGCATC ACTCAGACCA TGAGCGTCAC TCAGACCATG 480
15 AGCATCACTC AGACCAAGC CATCACTCTG ACCATGATCA TCACCTCCAC CATAATCATG 540
CTGCTTCTGG TAAAAATAAG CGAAAAGCTC TTTGCCGAGA CCATGACTCA GATAGTTCAG 600
GTAAAGATCC TAGAAACAGC CAGGGGAAAG GAGCTCACCG ACCAGAACAT GCCAGTGGTA 660
GAAGGAATGT CAAGGACAGT GTTAGTGCTA GTGAAGTGAC CTCACCTGTG TACAACACTG 720
TCTCTGAAGG AACTCACTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAATCTTTCC 780
20 CCAGAGATAT AAGCAGCTCC ACTCCACCCA GTGTCAATC AAAGAGCGCG GTGAGCGCGC 840
TGGCTGGTAG GAAACAAAT GAATCTGTGA GTGAGCGCGG AAAAGGCTTT ATGTATTCCA 900
GAAACACAAA TAGAAACAGC CAGGAGTGT TCAATGCATC AAAGCTACTG ACATCTCATG 960
GCATGGGCAT CCAGGTTCCG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA 1020
TCAACCAAT TGATGCTAGA TCTTGTCTGA TTCATACAAG TGAAAAGAAG GCTGAAATCC 1080
25 CTCAGAGAT CTATTCAAT CAAATAGCCT GGGTTGGTGG TTTTATAGCC ATTTTCATCA 1140
TCAGTTTCTC GTCTCTGCTG GGGGTATCT TAGTGCTCT CATGAATCGG GTGTTTTTCA 1200
AAITTCCTCT GAGTTTCTCT GTGGCACTGG CCGTTGGGAC TTTGAGTGGT GATGCTTTT 1260
TACACCTCTC TCCACATTCT CATGCAAGTC ACCACCATAG TCATAGCCAT GAAGAACCCAG 1320
CAATGGAAAT GAAAGAGGAA CCACCTTTCA GTCATCTGTC TTCTCAAAAC ATAGAAGAAA 1380
30 GTGCTTATT TGATTCCAGC TGGAGGGGTC TAACAGCTCT AGGAGGCGTG TATTTTATGT 1440
TTCTTGTGA ACATGCTCTC ACATTGATCA AACAATTAA AGATAAGAA AAAAAGAAATC 1500
AGAAGAAACC TGAAAATGAT GATGATGTGG AGATTAGAA GCAGTTGTCC AAGTATGAAT 1560
CTCACTTTT AACAAATGAG GAGAAAGTAG ATACAGATGA TCGAACTGAA GGCTATTTAC 1620
GAGCAGACTC ACAAGAGGCC TCCACCTTTG ATCTCAGCA GCCTGCAGTC TTGGAAGAAG 1680
35 AAGAGGTCTC GATAGTCTAT GCTCATCCAC AGGAAGTCTA CAATGAATAT GTACCCAGAG 1740
GGTGAAGAA TAAATGCCAT TCACATTTCC ACGATACACT CGGCCAGTCA GACGATCTCA 1800
TTCAACACCA TCATGACTAC CATCATATTC TCCATCATCA CCACCAACCA AACCAACATC 1860
CTCAGAGTCA CAGCAGCGCG TACTCTCGGG AGGAGCTGAA AGATGCGCGC GTCGCCACTT 1920
TGGCCTGGAT GGTGATAATG GGTGATGCGC TGCACAAAT CAGCGATGGC CTAGCAATTG 1980
40 GTGCTGCTTT TACTGAAGCG TTATCAAGTG GTTTAAGTAC TTCTGTGCT GTGTTCTGTC 2040
ATGAGTGGCC TCATGAATTA GGTGACTTTG CTGTTCTACT AAAGGCTGGC ATGACCGTTA 2100
AGCAGGCTGT CCTTATAAT CATTGTCTAG CCATGCTGGC GTATCTTGA ATGGCAACAG 2160
GAATTTTCTT TGTCATTAT GCTGAAATG TTTCTATGTG GATATTGCA CTTACTGCTG 2220
GCTTATTCTC GTATGTTGCT CTGTTTGATA TGGTACCTGA AATGCTGCAC AATGATGCTA 2280
45 GTGACGATAG GTATGCTGCG TGGGGGTATT TCTTTTACA GAATGCTGGG ATGCTTTTGG 2340
GTTTGTGAAT TATGTTACTT ATTTCCATAT TTGAACATAA AATCGTGTG CGTATAAAT 2400
TCTAGTTAAG GTTAAATATC TAGAGTAGCT TAAAAAGTTG TCATAGTTTC AGTAGGTCT 2460
AGGAGATGA GTTTGTATGC TGTACTATGC AGCGTTTAAA GTTAGTGGT TTTGTGATT 2520
TGTATTGAA TATGCTGTC TGTACAAAG TCAGTTAAAG GTACGTTTAA ATATTTAAGT 2580
50 TATTCTATCT TGGAGATAAA ATCTGTATGT GCAATTCACC GGTATTACCA GTTTATTATG 2640
TAAACAGAG ATTTGGCATT ACATGTTCTG TATGTTTCAG GGAATAATGT CTTTAAATGT 2700
TTTTCAAGAA CTAACACAGT TATTCCTATA CTGGATTTTA GGTCTCTGAA GAACCTGCTG 2760
TGTTTAGGAA TAAGAAATGT CATGAAGCCT AAAATACCAA GAAAGCTTAT ACTGAATTTA 2820
AGCAAGAGAA TAAAGAGAGAA AAGAGAAGAA TCTGAGAATT GGGGAGGCAT AGATTCTTAT 2880
55 AAAAATCACA AAATTTGTTG TAAATTAGAG GGGAGAAATT TAGAATTAAG TATAAAAGG 2940
CAGAATTAGT ATAGAGTACA TTCAATAAAC ATTTTGTGCA GGATTATTTC CCGTAAAAAC 3000
GTAGTGAGCA CTCTCAATGA CTAATTAGTG TACATTAAAC TTTGTATAAT ACAGAAATCT 3060
AAATATATTT AATGAATTCA AGCAATATAC ACTTGACCAA GAAATTGGAA TTTCAAAATG 3120
TTGTTGCGGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGTATCAC CAGACTGGGT 3180
60 TATTGCCAAG TTATATATCA CCAAAAGCTG TATGACTGGA TGTTCGTGTT ACCTGGTTTA 3240
CAAAATTATC AGAGTAGTAA AACTTTGATA TATATGAGGA TATTAATACT ACCTAAGTA 3300
TCATTGATT CGATTGAGAA AGTACTTGA TATCTCTCAG TGCTTCAGTG CTATCATTTG 3360
GAGCAATTGT CTTTATATAC GGTACTGTAG CCATACTAGG CCTGTCTGTG SCATTCTCTA 3420
65 GATGTTTCTT TTTTACACAA TAAATTCCTT ATATCAGCTT G
```

Seq ID NO: 415 Protein sequence
Protein Accession #: XP_084007

```
70 1 11 21 31 41 51
MARKLSVILI LTFALSVTNP LHELKAAAFP QTTEKISPWW ESGINVDLAI STRQYHLQQL 60
FYRYGENNSL SVEGFRKLQ NIGIDKIKRI HIHHDHDS DHEHSDHER HSDHEHSDH 120
EHHSDDHDS HHNHAASGN KRKALCPDHD SDSSGKDPN SQKGARPE HASGRNVD 180
SVSASEVTST VYNTVSEGT FLETIETPRP GKLFPPKDVSS STPPSVTSKS RVSRLAGRKT 240
75 NESVSEPRKG FMYSRNTNEN PQBCFNASKL LSHGMIQV PLNATEFNXL CPAIINQIDA 300
RSLIHTSEK KAEIPPKTY LQIAVGGFI AISIISFLSL LGVILVPLMN RVFFKPLSLP 360
LVALAVGTLS GDAFLHLPH SHASHSHS HEEPAMEMKR GPLFSLSSQ NIBESAYFDS 420
TWKGLTALGG LYFMLVHEV LTLIKQFKDK KKKNQKKPEN DDDVEIKKQL SKYESQLSTN 480
EEKVDDTDR EGYLRADSQE PSHFDSQQA VLEEESVMIA HAHPQEVNIE YVPRGCKNKC 540
80 HSHFDITLQQ SDDLHSHHD YHILEHHH QNHHPHSHS QYRSREELKDA GVATLAWMI 600
MGDGLNPSD GLAIGAFTE GLSSGLSTSV AVFCHPELPE LGDFAVLLKA GMTVKQAVLY 660
NALSAMLAYL GMATGIFIGH YAENVSMWIF ALTAGLFMYV ALVDMVPEML HNDASDHGCS 720
RWGYFFLQNA GMLLGGFIML LISIFEHKIV FRINP
```

85 Seq ID NO: 416 DNA sequence
Nucleic Acid Accession #: NM_015419.1
Coding sequence: 1..8487

	1	11	21	31	41	51	
5	ATGCCCAAGC	GCGCGCACTG	GGGGGCCCTC	TCCGTGGTGC	TGATCCTGCT	TTGGGGCCAT	60
	CCGCGAGTGG	CGCTGGCCTG	CCCGCATCCT	TGTGCTGCT	ACGTCCCCAG	CGAGSTCCAC	120
	TGCACGTTC	GATCCCTGGC	TTCCGTGCCC	GCTGGCATTG	CTAGACACGT	GGAAAGAAATC	180
	AATTITGGGT	TTAATAGCAT	ACAGGCCCTG	TCAGAAACCT	CATTITGCAGG	ACTGACCAAG	240
	TTGGAGCTAC	TTATGATTCA	CGGCAATGAG	ATCCCAAGCA	TCCCGATGG	AGCTTTAAGA	300
10	GACCTCAGCT	CTCTTCAGGT	TTTCAAGTTC	AGCTACAACA	AGCTGAGAGT	GATCACAGGA	360
	CAGACCCCTC	AGGGTCTCTC	TAACTTAATG	AGGCTGCACA	TTGACCACAA	CAAGATCGAG	420
	TTTATCCACC	CTCAAGCTTT	CAACGGCTTA	ACGTCTCTGA	GGCTACTCCA	TTTGGAAAGGA	480
	AATCTCCTCC	ACCAAGCTGCA	CCCCAGCACC	TTCTCCACGT	TCACATTTT	GGATTATTTT	540
	AGACTCTCCA	CCATAAGGCA	CCTCTACTTA	GCAGAGAACA	TGGTTAGAAC	TCTTCTGCC	600
15	AGCATGCTTC	GGACATGCC	GCTTCTGGAG	AATCTTTACT	TGCAGGGAAA	TCCGTGGACC	660
	TGCGATTGTG	AGATGAGATG	GTTTTGGAA	TGGGATGCAA	AATCCAGAGG	AATTCTGAAG	720
	TGTAAGAAAG	ACAAAGCTTA	TGAAGCGGT	CAGTTGTGTG	CAATGTGCTT	CAGTCCAAAG	780
	AAGTTGTACA	AACATGAGAT	ACACAAGCTG	AAGGACATGA	CTTGTCTGAA	GCCTTCAATA	840
	GAGTCCCTCT	TGAGACAGAA	CAGGAGCAGG	AGTATTGAGG	AGGAGCAAGA	ACAGGAAGAG	900
20	GATGGTGGCA	GCCAGCTCAT	CCTGGAGAAA	TTCCAATGTC	CCCAGTGGAG	CATCTCTTTG	960
	AATATGACCG	ACAGTCTGGA	GAAATCTGGT	AACTTGGTCT	GTGACATCAA	GAAGCAAAATG	1020
	GATGTGTAC	AGATTCACTT	GAACCAAAAG	GATCCTCCAG	ATATTGACAT	AAATGCAACA	1080
	GTTCCTTGG	ACTTTGAGTG	TCCAATGACC	CGAGAAAATC	ATGAAAAGCT	ATGGAATTTG	1140
	ATAGCATACT	ACATGAAGT	TCCCGTGAAG	CTACACAGAG	AGCTCATGCT	CAGCAAAAGAC	1200
25	CCCAGAGTCA	GCTACAGTGA	CAGGACAGAT	GCTGATGAGG	AAGCTCTTTA	CTACACAGGT	1260
	GTGAGAGCCC	AGATTCTTGC	AGAACCCAGAA	TGGGTCTGTC	AGCCATCCAT	AGATATCCAG	1320
	CTGAACCGAC	GTCAAGATAC	GGCCAAGAA	GTGCTACTTT	CCTACTACAC	CCAGTATTCT	1380
	CAAAATATAT	CCACCAAGAA	TACAAGGCAG	GCTCGGGGCA	GAAGCTGGGT	AATGATTGAG	1440
	CCTAGTGGAG	CTGTGCAAG	AGATCAGACT	GTCCCTGGAAG	GGGTCCATG	CCAGTTGAGC	1500
30	TGCAACGTGA	AAGCTTCTGA	GAGTCCATCT	ATCTTCTGGG	TGCTTCCAGA	TGGCTCCATC	1560
	CTGAAAGCGC	CATGGGATGA	CCGAGACAGC	AAGTTCTCCA	TTCTCAGCAG	TGGCTGGCTG	1620
	AGGATCAAGT	CCATGGAGCC	ATCTGACTCA	GGCTTGTACC	AGTGCAITGC	TCAAGTGAAG	1680
	GATGAAATGG	ACCCCATGGT	ATATAGGGA	CTTGTGCAGT	CTCCCTCCAC	TCAGCCAGCC	1740
	GAGAAAGACA	CAGTGACAAT	TGGCAAGAAC	CCAGGGGAGT	CGGTGACATT	GCCTTGCAAT	1800
35	GCTTTAGCAA	TACCCGAAGC	CAACCTTAGC	TGGATTCTTC	CAACAGAAAG	GATAATTAAT	1860
	GATTTGGCTA	ACCATACACA	TGTATACATG	TTGCCAAATG	GAATCTTTTC	CATCCCAAG	1920
	GTCCAAGTCA	GTGATAGTGG	TTACTACAGA	TGTGTGGCTG	TCAACCAGCA	AGGGGCAGAC	1980
	CATTTTACCG	TGGGAATCAC	AGTGACCAAG	AAAGGGTCTG	GCTTGCCATC	CAAAAGAGGC	2040
40	AGAGCCCTAG	GTGCAAAAGC	TCTTTCCAGA	GTCAAGAGAG	ACATCGTGG	GGATGAAGGG	2100
	GGCTCGGGCA	TGGGAGATGA	AGAGAACACT	TCAAGGAGAC	TTCTGCATCC	AAAGGACCA	2160
	GAGGTGTTC	TCAAAACAA	GGATGATGCC	ATCAATGGAG	ACAAGAAAGC	CAAGAAAGGG	2220
	AGAAGAAAGC	TGAAACTCTG	GAAGCATTCG	GAAGAAAGAC	CAGAGACCA	TGTTGCAGAA	2280
	GGTCCGAGAG	TGTTTGAATC	TAGACGAAGG	ATAAACATGG	CAAAACAAAC	GATTAAATCC	2340
45	GAGCGCTGGG	CTGATATTTT	AGCCAAAGTC	CGTGGGAAAT	ATCTCCTTAA	GGGCACAGAA	2400
	GTACCCCGCT	TGATTAAAC	CACAGTCTCT	CCATCCTTGA	GCCTAGAAGT	CACACCACT	2460
	TTTCTCCTG	TCTTCCCTC	CTCAGCATCT	CCTGTGCAGA	CAGTAACCA	TGCTGAAGAA	2520
	TCCTCAGCAG	ATGTACCTCT	ACTTGGTGAA	GAAGAGCACG	TTTTGGGTAC	CATTCTCTCA	2580
	GCCAGCATGG	GGCTAGAAC	CAACCACAAT	GGAGTTATTC	TTGTTGAACC	TGAAGTAACA	2640
50	AGCACACCTC	TGGAGGAAGT	TGTTGATGAC	CTTCTGAGA	AGACTGAGGA	GATAACTTCC	2700
	ACTGAAGGAG	ACCTGAAGGG	GACAGCAGCC	CCTACACTTA	TATCTGAGCC	TTATGAACCA	2760
	TCTCTACTCT	TGCACACTT	AGACACAGTC	TATGAAAGC	CCACCATGGA	AGAGACGGCA	2820
	ACAGAGGGTT	GGTCTGCAGC	AGATGTTGGA	TGTCACCAAG	AGCCACATC	CAGTGAGTAT	2880
	GAGCCTCCAT	TGGATGCTGT	CTCCTTGGCT	GAGTCTGAGC	CCATGCAATA	CTTTGACCCA	2940
55	GATTTGGAGA	CTAAGTCACA	ACCAGATGAG	GATAAGATGA	AAGAAGACAC	CTTTGCACAC	3000
	CTTACTCCAA	CCCCACCAT	CTGGGTAAAT	GACTCCAGTA	CATCAGAGT	ATTGAGGAT	3060
	TCTACTTAG	GGGAACAGG	TGTCCAGGC	CAATCACATC	TACAAGGACT	GACAGACAAC	3120
	ATCCACCTTG	TGAAAGTAG	TCTAAGCACT	CAAGACACTT	TACTGATTAA	AAAGGGTATG	3180
	AAAGAGATGT	CTCAGACACT	ACAGGGAGGA	AAATGCTAG	AGGAGACCC	CACACACTCC	3240
60	AGAAGTTCTG	AGAGTGAGGG	CCAAGAGAGC	AAATCCATCA	CTTTGCTGGA	CTCCCACTG	3300
	GGTATAATGA	GCAGTATGTC	TCCAGTTAAG	AAGCCTCGGG	AAACCAAGT	TGGTACCCTC	3360
	CTAGACAAG	ACACCAACA	AGTAACAACA	ACACCAAGGC	AAAAAGTTGC	TCCGTCTATC	3420
	ACCATGAGCA	CTCACCTTTC	TGGAAGGAGA	CCCAACGGGA	GAAGGAGATT	ACGCCCAAC	3480
	AAATTCCGCC	ACCGGCACAA	GCAAAACCCA	CCCAACAATT	TTGCCCATC	AGAGACTTTT	3540
65	TCTACTCAAC	CAACTCAAGC	ACCTGACATT	AAGATTTCAA	GTCAAGTGG	GAGTCTCTG	3600
	GTTCTACAG	CTGGGTGGA	TAAACAGATT	AATACCCCA	AACAGTTGGA	AATGGAGAAG	3660
	AATGCAGAG	CCACATCCAA	GGGAACACCA	CGGAGAAAGC	ACGGGAAGAG	GCCAAACAA	3720
	CATCGATATA	CCCCTTCTAC	AGTGAGCTCA	AGAGCGTCCG	GATCCAAGCC	CAGCCCTTCT	3780
	CCAGAAAATA	AACATAGAAA	CATTGTTACT	CCCAGTTTCA	AACTATACT	TTTGCTTAGA	3840
70	ACTGTTTCTC	TGAAAACCTGA	GGGCCCTTAT	GATTCTTAG	ATTACATGAC	AACCACCAGA	3900
	AAAAATATAT	CATCTTACCC	TAAAGTCCAA	GAGACACTTC	CAGTCACATA	TAAACCCACA	3960
	TCAGATGGAA	AAGAAATTA	GGATGATGTT	GCCACAAATG	TTGACAAACA	TAAAGTGAC	4020
	ATTTTAGTCA	CTGGTGAATC	AATTACTAAT	GCCATACCAA	CTTCTGCTC	CTTGGTCTCC	4080
	ACTATGGGAG	AATTTAAGGA	AGAATCTCT	CCTGTAGGCT	TTCCAGGAAC	TCCAACCTGG	4140
75	AATCCCTCAA	GGACGGCCCA	GCCTGGGAGG	CTACAGACAG	ACATACCTGT	TACCATTCT	4200
	GGGGAATAAT	TTACAGACCC	TCCCTTCTT	AAAGAGCTTG	AGGATGTGGA	TTTCACTTCC	4260
	GAGTTTGTG	AGCTTTTGTG	AGTCTCCACA	CCATTCTACC	AGGAAGAAGC	TGGTTCTTCC	4320
	ACAACCTCTCT	CAAGCATAAA	AGTGGAGGTG	GCTTCAAGTC	AGGCAGAAAC	CACCACCTTT	4380
	GATCAAGATC	ATCTTGAAC	CAGTGTGGCT	ATTCTCTCTT	CTGAAACTAG	ACCACAGAA	4440
80	CACACCCCTA	CTGCTGCCCG	GATGAAGGAG	CCAGCATCTC	CGTCCCATC	CACAAATCTC	4500
	ATGTCCTTGG	GACAAACCA	CACCACTAAG	CCAGCACTTC	CCAGTCCAAG	AATATCTCAA	4560
	GCATCTAGAG	ATTCCAAGGA	AAATGTTTTC	TTGAATTATG	TGGGAATCC	AGAAACAGAA	4620
	GCAACCTCAG	CAACAATGA	AGGAACACAG	CATATGTCAG	GGCCAAATGA	ATTATCAACA	4680
	CCCTCTTCCG	ACCGGATGTC	ATTTAATCTG	TCTACAAAGC	TGGAATTGGA	AAAGCAAGTA	4740
	TTTGGTAGTA	GGAGTCTTACC	ACGTGGCCCA	GATAGCCAA	GCCAGGATGG	AAGAGTTTCT	4800
85	GCTTCTCATC	AACTAACCA	AGTCCCTGCC	AAACCCATCC	TACCAACAGC	AACAGTGAAG	4860
	CTACCTGAAA	TGTCCACACA	AAGCGCTTCC	AGATACTTTG	TAACTTCCCA	GTCACTCTGT	4920
	CACTGGACCA	ACAAACCGGA	AATAACTACA	TATCCTTCTG	GGGCTTTGCC	AGAGAACAAA	4980
	CAGTTTACAA	CTCCAAGATT	ATCAAGTACA	ACAATTCTCT	TCCATTGCA	CATGTCCAAA	5040

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80

CCCAGCAATC CTAGTAAGTT TACTGACOGA AGAACTGACC AATTCATGTT TTAATCCAAA 5100
 GTGTTTGGAA ATAACAACAT CCTGAGGCA AGAAACCCAG TTGGAAAGCC TCCCAGTCCA 5160
 AGAATTCCTC ATTATTCCAA TGGAGAGCTC CCTTTCTTTA CCAACAAGAC TCTTTCTTTT 5220
 CCACAGTTGG GAGTCACCCG GAGACCCAG ATACCCACTT CTCCTGCCCC AGTAATGAGA 5280
 GAGAGAAAAG TTATTCACAG TTCTTACAAC AGGATACATT CCCATAGCAC CTTCATCTCG 5340
 GACTTTGGCC CTCCGGCACC TCCGTTGTTG CACACTCCCG AGACCCAGGG ATCACCTCTA 5400
 ACTAATCTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC CTTTATAACA 5460
 TCTTCTGTCC AGTCTCAGG AAGCTTCCAC CAGAGCAGCT CAAAGTTCTT TGCAGGAGGA 5520
 CCTCTGCTAT CCAAAATCTG GTCTCTTGGG GAAAAGCCCC AAATCCTCAC CAAGTCCCCA 5580
 CAGACTGTGT CCGTCAACCG TGAGACAGAC ACTGTGTTCG CCTGTGAGGC AACAGGAAAA 5640
 CCAAAAGCCTT TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC 5700
 AGGATACAAC GGTTTGAGTT TCTCAAGAAC GGTACCTTAG TGATACGGAA GGTTCAGTA 5760
 CAAGATCGAG GCCAGTATAT GTGCACCCGC AGCAACCTGC ACGGCCTGGA CAGGATGTTG 5820
 GTCTTGCTTT CGGTCAACCG GCAGCAACCT CAAATCTAG CCTCCACTA CCAGGACGTC 5880
 ACTGTCTACC TGGGAGACAC CATTGCAATG GAGTGTCTGG CCAAGAGGAC CCGAGCCCCC 5940
 CAAATTTCTT GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC CGTGGAGAGC 6000
 CGCATCACCC TGACGAAAAA CCGGACCTTT TCCATCAAGG AGGGCTCTCT CTCAGACAGA 6060
 GGGGTCTATA AGTGGTGGG CAGCAATGCA GCGGGGGCGG ACAGCCTGGC CATCCGCTCT 6120
 CAGTGGGGCG CACTGCCCCC CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC 6180
 CCGGGGCTCA CGGTACACAT TCACGTGACT GCCAAGGCTG CCGCCCTGCC CAGCGTGGCG 6240
 TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCTCTGCACT TCCTTCAAGG GAACTTGTCT 6300
 GTTTTCCCAA ACGGACGCT CTACATCCGC AACCTGCGC CCAAGGACAG GCGGGCTAT 6360
 GAGTGGGTGG CCGCCAAACCT GGTAGGCTCC GCGCGCAGGA CGGTGACGCT GAACTGTCAG 6420
 CGTGCAGCAG CCAACGCGCG CATCACGGGC ACCTCCCGCG GGAGGACGGA CGTCAGGTAC 6480
 GGAGGAACCC TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCGCGC CATCCTCTGG 6540
 AGGCTCGCCT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG AATCAAGGTG 6600
 TTTGCCAATG GACCCCTGGT GGTGAAATCA GTGACGGACA AAGATGCGCG AGATTACCTG 6660
 TGGGTAGCTC GAAATAAGGT TGGTGAATGAC TAAGTGTGTC TCAAAGTGGG TGTGGTGATG 6720
 AAACCGGCCA AGATTGAACA CAAGGAGGAG AACGACCACA AAGTCTTCTA CCGGGGTGAC 6780
 CTGAAAGTGG ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCTCG GAGCCTCCCA 6840
 GACGGGAGTC TGGTGAATCT CTTCATGACG TCGGATGACA GCGGTGGAGC CACCAAGGCG 6900
 TATGTCTGCT TCAACAATGG GACACTCTAC TTTAACGAAG TGGGGATGAG GGAGGAAGGA 6960
 GACTACACCT GCTTTGCTGA AATCAGGTG GGAAGGAGC AGATGAGAGT CAGAGTCAAG 7020
 GTGGTGACAG CGCCCGCCAC CATCCGGAAC AAGACTTACT TGGCGTTTCA GGTGCCCTAT 7080
 GGAGACGTGG TCAGTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG 7140
 TGTTCGCCAA CCAACAAGGT GATCCCGACC TCCTCTGAGA AGTATCAGAT ATACCAAGAT 7200
 GGCACTCTCC TTATTCAGAA AGCCGACGCT TCTGACAGCG GCAACTACAC CTGCTGTGTC 7260
 AGGAACAGCG CGGGAGAGGA TAGGAAGAGC GTGTGGATTG ACGTCAACGT CCAGCCACCC 7320
 AAGATCAACG GTAACCCCAA CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGCGAGT 7380
 CGGAAACTGA TTGACTGCAA AGCTGAAGGC ATCCCAACCC CGAGGGTGTG ATGGGCTTTT 7440
 CCGGAGGTGG TGGTCTGCCC AGCTCCATAC TATGAAACCC GGATCACTGT CCATGGCAAC 7500
 GGTTCCTGGT ACATCAGGAG TTTGAGGAAG AGCGACTCCG TCCAGCTGGT ATGCTAGGCA 7560
 CGCAACGAGG GAGGGGAGGC GAGGTTGATC GTGCAGTCA CTGTCTGGA GCCCATGGAG 7620
 AAACCCATCT TCCACGACCC GATCAGCGAG AAGATCAACG CCATGCGCGG CCACACCATC 7680
 AGCTTCAACT GCTCTGAGCG GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT 7740
 GGCAACGATC TGACAGAGTG ACAGCAGCTG CAGCGCTTCT ACCACAAGGC TGACGGCATG 7800
 CTACACATCT TGAGTCTTCT CTGCTGGGAC GCTGGGGGCT ACCGCTGCGT GGGCCGCAAT 7860
 GCGCTGTGCC ACACGAGAGG GCTGTCTCC CTGAAGGTGG GACTGAAGCC AGAAGCAAAC 7920
 AAGCAGTATC ATAACCTGGT CAGCATCATC AATGTTGAGA CCTGAAGCT CCCCTGCACC 7980
 CCTCCCGGGG TGGGCGAGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG 8040
 GGCCCCCAAA CCTTGGGAGC CGTTTCTCTT CTGGAACAAT GCACCTTCAC GGTTCGTGAG 8100
 GCGCTCGGTG TTGACAGGGG TACCTATGTA TGCAAGATGG AGACGGAGTA CGGCCCTTCG 8160
 GTACCCAGCA TCCCGTGATG TGTGATCGCC TATCCTCCCC GGATCACCAG CGAGCCACCC 8220
 CCGGTCACTT ACACCCGCGC CGGGAACACC GTGAACTGA ACTGCATGGC TATGGGGATT 8280
 CCAAGAGCTG ACATCACTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTCAG 8340
 GCTGCTGTGT ATGGAAACAG ATTTCTTCAC CCCCAGGGAT CACTGACCAT CCAGCATGCC 8400
 ACACAGAGAG ATGCGCGGCT CTACAAGTGC ATGGCAAAAA ACATTCTCGG CAGTGACTCC 8460
 AAAACAACCT ACATCCAGT CTCTGAAAT GTGGATTCCA GAATGATTGC TTAGGAACATG 8520
 ACAACAAAGC GGGGTTTGTG AGGGAAGCCA GGTGGGGGAA TAGGAGCTCT TAAATAATGT 8580
 GTCACAGTGC ATGGTGGGCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT 8640
 GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT GAGACACTTT 8700
 CTTTGTGTGT TACATCATGC CAGGGGCTTC ATTGAGGGTG TCTGTGCTCT GACTGCAATT 8760
 TTTCTTCTTT TGCAATGCC ACTCGACTGC CTTCATAAGC GTCCATAGGA TATCTGAGGA 8820
 ACATTATGTA AAAATAAGCC ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA 8880
 TCACCTAGTT AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCAGAT TGACAAGTCA 8940
 TCTTTCAAGT ATTCTCTCTG TCACITCAAAC ACTCCAGCTT GCCCAATAAG GATTTAGAAC 9000
 CAGAGTGACT GATATATATA TATATATTTT AATTCAAGT TACATACATA CAGCTACCAT 9060
 TTTATATGAA AAAAGAAAAA CATTTCTTCC TGGAACTCAC TTTTATATA ATGTTTATA 9120
 TATATATTTT TTCTTTCAA ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT 9180
 ATTAAATTA ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA 9240
 ATATAATTTT AAAAATTTC TCTCCAACCT CCTTCAAAAT CAGTCACCACT TGTATATTA 9300
 CCTTCTCCAG GAACCTCCA GTGGGGAAGG CTGCGATATT AGATTTCCTT GTATGCAAG 9360
 TTTTGTGTA AAGCTGTGCT CAGAGGAGGT GAGAGGAGAG GAAGGAGAAA ACTGCAATAT 9420
 AACTTTACAG AATTGAATCT AGAGTCTTCC CCGAAAGGCC CAGAACTTC TCTGCAGTAT 9480
 CTGGCTGTGC CATCTGTGCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC 9540
 CCATGAATAA TACACGACCT GTTATTCCA TGACTGCTTT ACTGTATTTT TAAGGTCAAT 9600
 ATACTGTATA TTTGATAATA AAATAATATT TCCCAAAAA AAAAA

Seq ID NO: 417 Protein sequence
 Protein Accession #: NP_056234.1

1 11 21 31 41 51
 85 MPKRAHWGAL SVVLLILLWGH PRVALACPHP CACYVPSEVH CTFRSLASVP AGIARHVERI 60
 NLGFNSIQAL SETSPAGLTK LELLMIHNE IPSIPDGLR DLSSQVFKP SYNKLRVITG 120
 QTLQGLSNLM RLHIDHNKIE FIHPQAFNGL TSLRLHLEG NLLHLQHPST FSTFTFLDYF 180
 RLSTIRHLYL AENMYRTLPA SMLRNMPLE NLYLQGNFWT CDCEMRWFLS WDAKSRGILK 240

CKKDKAYEGG QLCAMCFSPK KLYKHEIHLK KDMTCLKPSI ESPLRQNRSR SIEEDQEQEE 300
DGGSQLILEK FQLPQWSISL NMTDEHGNMV NLVCDIKKPM DVYKIHNLQT DPPDIDINAT 360
VALDFECPMY RENYEKLWLK IAYYSEVPVK LHRELMLSKD PRVSYQYRQD ADEEALYYTG 420
VRAQILAEPE WVMQPSIDIQ LNRQSTAKK VLLSYTYQYS QTISTKTRQ ARGRSWVMIE 480
PSGAVQRQDT VLEGGPCQLS CNVKASESPS IFWVLPDGI LKAFMDPDPS KPSILSSGWL 540
RIKSMEPSDS GLYQCIQAVR DEMDRMVYRV LVQSPSTQPA EKDTVTIGKN PGESVTLPCN 600
ALAIPEAHLN WILFNRIIN DLANTSHVYM LPNGTSLIPK VQVSDSGYR CVAVNOQCAD 660
HFTVGTITVK KSGGLPSKRG RRPKAKALSR VREDIVEDG GSGMGDEENT SRRLHHPKQD 720
EVLKTKDDA INGDKKAKKG RRLKLNKHS EKEPETNVAE GRVVFESRRR INMANKQINP 780
ERWADILAKV RGNLPLKGTG VPPLIKTTSP PSLSLEVTPP FPAVSPSPAS PVQVTSAEE 840
SSADVPLLE EHVLTGISS ASMGLEHNHN GVILVEPEVT STPLEEVDD LSEKTEEITS 900
TEGDLKGTAA PTLISEPYEP SPTLHTLDTV YEKPTHEETA TEGWSAADVG SSSEPTSSEY 960
EPPLDAVSLA ESEPMQYFDP DLETKSQPDE DKMKEDTFAH LTPTPTIWNV DSSTSOLFED 1020
STIGEPVPG QSHLQGLTDN IHLVKSSLSL QDTLLIKKGM KEMSTLQGG NMLEGDPHVS 1080
RSSESEGGDS KSITLPSDTL GIMSSMSPVK KPAETTVGTL LDKDITTVTT TPRQKVAPSS 1140
TMSHPSRRR PNGRRRLRKH KFRHRHKQTP PTTTAPSETF STQPTQAPDI KISSQVSESL 1200
VPTAWVDNTV NTFKQLEMEK NAEPSTKGT PRRKHGKRPNK HRYTPSTVSS RASGSKPSPS 1260
PENKRNIVT PSSETILLPR TVSLKTEGYP DSLDYMTTTR KIYSSYPKVQ ETLFVITYKPT 1320
SDGKETKDDV ATNVDKHKS ILVTGESITN AIPSTRSLVS TMGEFKEESS PVGPPGTPTW 1380
NPSRTAQPR LQDTLPVTS GENLTDPLL KELEDVDFTS EFLSSTVST PFHQEAGSS 1440
TTLSSIKVEV ASSQAEITTL DQDHELTVA ILLSETRPQN HTPTAARMKE PASSSPSTIL 1500
MSLGQTITTK PALPSRISQ ASRDSKENVF LNYVGNPETE ATFVNNEGTO HMSGPNELST 1560
PSSDRDAFNL STKLELEKQV FGSRSLPRGP DSQRQDGRVH ASHQLTRVPA KPILPTATVR 1620
LPPMSTQAS RYFVTSQSPR HWTNKPEITT YPSGALPENK QFTTPRLST TIFPLHMSK 1680
PSIPSKPTDR RTQDIPVYSK VFGNNIPEA RNPVGKPPSP RHPHYNGLR PFTNKTLSF 1740
PQLGVTRRPQ IPTSPAPVMR ERKVIPGSYN RIHSHSTFHL DFGPPAPPLL HTPQTGSPS 1800
TNLQNPVMS STQSSISFIT SSVQSSGSEH QSSSKFPAGG PPASKFWSLG EKPOILTSP 1860
QTVSVTAETD TVPCEATGK PKPFVTVTKV STGALMTPNT RIQFEVLKN GTLVIKRVQV 1920
QDRQYMCCTA SNLHGLDRMV VLLSVTVQQP QILASHYQDV TVYLGDTIAM ECLAKGTAP 1980
QISWIFPDR VWTQSPVYS RITLHENRTL SIKEASFSDR GYVKCVASNA AGADSLAIRL 2040
HVAALPPVH QEKLENIPL PGLSIHICT AKAPLPSVR VWLGDGTQIR PSQFLHGNLF 2100
VFPNGTYLR NLAPKDSGRY ECVAANLVGS ARRTVQLNVQ RAAANARITG TSPRRTDVRY 2160
GGTLKLDCA SGPDPFRILW RLPSKRMIDA LFSFDSRIK FANGTLVVK VTDKAGDYL 2220
CVARNKQDD YVVLKVDVVM KPAKIEHKEE NDHKVFYGGD LKVDVCATGL PNPEISWSLP 2280
DGLSVNFMQ SDDSGRTRK YVFPNNGTLY FNEVGMREEG DYTCAENQV GKDEMVRVRK 2340
VVTATPATRN KTYLAVQVPE GDVVTVACEA KGEPMKVTW LSPTNKVIPT SSEKYIYQD 2400
GTLLIQAKR SDSGNYTCLV RNSAGEDRKT VWIHNVPFP KINGNPNPT TVREIAAGGS 2460
RKLIDCKAEG IPTPRVLWAF PEGVVLPAFY YGNRITVHGN GSLLDIRSLR SDSVQLVCA 2520
RNEGGEARLI VQLTVLEPME KPFIHDPICE KITAMAGHTI SLNCSAAGTP TPLSVVWLPN 2580
GTDLQSGQL QRFYHKAQDM LHSGLSSVD AGAYRCVARN AAGHTERLVS LKVGKPEAN 2640
KQYHNLVSI NGETLKLCT PPGAGQGRFS WTLFNGMHLE GPQTLGRVSL LDNGTLTVRE 2700
ASVDFRGTIV CRMETEYGPS VTSIPVIVIA YPRITSEPT FVIYTRPGNT VKLNCMAMGI 2760
PKADITWELP DKSHLKAGV ARLYGNRFLH PQGSLTIQHA TQRDAGFYK MAKVILGSDS 2820
KTYIHVF

Seq ID NO: 418 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1...5001

1 11 21 31 41 51
ATGCCAGGCA CAAACTAAC CGAACAGCG GCCCAGCAG ACTACAGAGT GATATTGAAG 60
ACCTCTCAAG AGGACGAATT GGATGTACCT GACGACATCA CGCTCCGGGT TATGTTCATCT 120
CAGTCTGTGC TTGTGTCTCG GTTGATCCT GTTCTGGAAA AACAGAAGAA AGTTGTTGCA 180
TCAAGACAGT ACACGCTGCG CTATCGAGAG AAGGGGGAAT TGGCCAGGTG GGATTATAAG 240
CAGATCGCTA ACAGCGCTGT GCTGATTGAG AACCTGATTC CAGACACTGT GTATGAATTT 300
GCATCCGTA TTTCACAGG TGAAAGAGAT GCCTAATGGA GTACGTCACT CTTCACAAAG 360
ACACCAGAA: CTGCCCTAC CACAGCTCCT GAAACTTGA ACGTCTGGCC AGTCAATGGC 420
AAACCTACAG TTGTGCTGTC ATCTTGGGAT GCGCTACCAG AGACTGAGG GAAAGTGAAA 480
GTCTGTCTGC TGGACACAGG ACTGTTTCA GTTCTCTCCT TCCAAACATC TGCCAAATCA 540
TTTCAGAATA CATCTTTCA TACGCCCGG CTCTCAAACC ATTTGGAGCA AAGTCCCTCA 600
CCTATCTCG: AGACACTACT TCTGCCCTGG TGGATGGTCT GCAGCCTGGG GAACGCTATC 660
TTTTCAAAAT: CCGGGCCACA AACAGGAGAG GCCTGGGACC TCACTCCAAA GCCTTCATG 720
TCGCTATGCC AACAGAATG CAGCTGTACC CAGAAGGATT TCAGTTGTCT AGCTTACCTG 780
ATCGATATCC AAACCAACA AGTTAATAAA GATCCACAAC TGGAGGGAGT TGTTTTGGAA 840
CCATGTTTTT TTTTCTACTT CCTCACATT ATGCTGGATA TTGGCGGCTT TTCTTCATT 900
ATGTGCTATG AAGACCCANN TGTTCTTCT TTGACAGGCA ATTCTTAAA ATCTGTTGCA 960
GCCAGTAAGG CGGATGTTCA GCAGAACAGG GAGGACATG GGAACCCGA AAAACCTGAG 1020
CCTTCTCAC CTCTCCAG AGCTCCAGCT TCCTCCCAAC ACCCTCTGT GCCTGCTTCT 1080
CCCCAAGGGA: GAAATGCCAA GGACCTTCTT CTGACTTGA AGAACAAAAT ATTGGCTAAT 1140
GGTGGGGCGC CCGGAAAACC CAGCTTCGC GCCAAGAGG CAGAGGAGCT GGATCTTCAG 1200
TCGACAGAAA TCACCTGGGA GAGGAGCTG GGTTCGGG AGGACTCGCC CATGTCAACC 1260
TCAGACACCC AAGACAGAA ACGGACCTG AGGCCGCCAA GTAGACAGG CCACTCGGTG 1320
GTTGCTCCCG GCAGGACTGC AGTGAGGCC CGGATGCCAG CGCTGCCCG AAGGGAAGGC 1380
GTAGATAAGC CTGGCTTTTC CTTGGCCAG CAGCCCGGCC CAGGGCGGCC CCCCTCGGCT 1440
TCGGCTCTC CTGCCACCA CGCTCCACC CAGGGCACCT CTCATCGTTC TTCCCTGCCT 1500
GCCAGCTTGA ATGACAAGA CTGTGTGAC TCAGACGAAG ATGAGCGCG TGTGGGCTCC 1560
CTCCACCCCA AGGGGCGCTT: CGCCAGGCC CGGCCAGGCC TGTCCCGG CCGCCAGTCC 1620
CGCTCCAGCG TTCTCCGGA CAGAAGCTCT GTGACCCCG GCGCAAGGCC AGCTCGGCC 1680
CGCGGAGGA CCCCCATT CAGGGGCCGA GAGGAAGATT CCACTGCTTC AGCCCAACC 1740
TCAGACTTTC TTCCACCCA TGGGGGATCA TCTCGGCTGC TGCCACCCA GCCACACCTG 1800
AGCTCTCCAC TTTCAGGG CGGGAAGGAT GGTGAGGAC CCGGAGCAC CAACTCCAAT 1860
GGCCCATCAC GGTCCACCAT GTCTCTCTCT ATCTCTCTCT CAGGACGCG 1920
GTCTCTGAGG GAGCGGAGGC TTCTGATGTT GAAAGCCAG GTGACGCGGA TAGGGAAGAC 1980
GGCGGAAGGC AGCGGAGGC CAGCGGCCAG ACCTGCGGG CCGGCGCTGC CTGGAACAC 2040
TTCCATTTCG TCAGACCAA ACCCTTGTCT GCCAAGGGA GGTCTCCAAG CAGGTTCAAG 2100
ATTGGGCGGG GACCTCGGCT GCAGCCCTCC AGCTCCCAAC AGTCACTGT GCGCTCCGA 2160

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

```

GCCACCCCA GGGTTCCCTC TCACTCTGAT TCCACCCCTA AGCTTAGCTC AGGTATCCAT 2220
GGAGACGAGG AGGATGAGAA GCGCTTCCTT GCCACCGTTG TCAATGACCA CGTGCCCTTCC 2280
TCCTCCAGGC AGCCCATCTC CCGGGGCTGG GAGGACTTAA GGAGAAGCCC GCAGAGAGGG 2340
GCCAGCCTCG ATCGGAAGGA ACCCATCCCA GAGAACCCCA AATCCACAGG GGCAGATACA 2400
CATCTCAGG GCAAGTACTC CTCCCTGGCC TCCAAAGGCTC AGGATGTTCA ACAGAGCACA 2460
GACGCGGACA CGGAGGCTCA TTCTCCCAAA GCACAGCCAG GGTCCACAGA CCGCAGCGCG 2520
TCCCTGCTCT GTCTCTCCGC AGCAGCGTCA CAGCAGCATC CCAGTGTTC CAGAAGGATG 2580
ACACCCGCGC GGGCCCCAGA ACAGCAGCCC CCTCTCCCG TCGCCACGTC CCAGCACCAC 2640
CAGGACCCCG AGACGAGAGA CGCGGCTGG TCACTTCCC AGCCACGGCT CTCACGACC 2700
CAGGCGCGGC GGGCCCCGCC CACGTCGCG GCGCGCTCCC ACTCTCTCTC GGACCCCTTAC 2760
AOCGCGAGCT CCAGAGGGAT GCTCCCCACG GCGCTCCAGA ACCAGGACGA GGATGCCCGAG 2820
GGCAGCTACG ACAGCAGACG CACAGAAATC GAGGCCCAGG ATGTGCGGGC CCGCGCGCAC 2880
GCGCGCGCGC CCAAGGAGGC AGCTGCGTCC CTTCACAGC ACCAGCAGGT GGAGTCTCCC 2940
ACAGSCCAGG GGGCAGTGG GACCCACAGG TCCAGCGCG GACATGCGGC CTCCCCGCC 3000
AGGCCAGGCC GAGCCGCGCG CCCCCAGTCC CGCGCCCGGG TCCCCAGCAG GGCAGCGCGC 3060
GGGAAGTCGG AGCTCTCTTC CAAGCGCGCC CTGTCTCCCA AGTCCACAGA GTCCGTCTCA 3120
GCGGAGGATC AGGAGGAGA GGAACGCGGG TTTTAAAG GCGGAAAGA AGACCTTCTG 3180
TCTTCTCTG TGCCAAAGTG GCGCTCTTCC TCCACTCCCA GGGCGGCAAG AGAGCGCGAT 3240
GGGAGCCTCG CCAAGGAAGA GAGGGAGCCT GCCATGCGCG TTGCGCTCG CGGAGGGAGC 3300
CTGGCTCTCT TGAGAGGACC TCTCCCCCA CCTCCAGGCA GCTCCCCCAG GGCCTCCAC 3360
GTCCCTTCCC GACCGCGCC TCGCAGCGCT GCCACCGTGA GCGCGCTGCG GGCACCCAC 3420
CCTTGGCGCG GTTACACCA CCGCGCCCCC CTGGCCACT TCTCCACCA CCGGATGCTG 3480
TCCTTGGCGC AGAGGATGAT GCATGCCAGA TTCCGTAAAC CTCTCTCCCG ACAGCCTGCC 3540
AGACCCCTTT ACAGACAAGG TTATAATGGC AGACCAAATG TAGAAGGGA AGTCCTTCTC 3600
GGTAGTAATG GAAACCGGAA TGGACAGAGA ATTATCAATG GCCCTCAAGG AACAAAGTGG 3660
GTTGTGGACC TTGATCGTGG GTTAGTATTG AATGCAGAAG GAAGGTACCT CCAAGATTCA 3720
CATGGAATCG CTCTTCGGAT TAACTAGGA GGAGATGGTC GAACCATGTG AGATCTGGAA 3780
GGGACCCCGC TGGTAGTCC TGACGGCTCT CCACTCTTTG GGCAGGGCGC ACATGGCACA 3840
CCTCTGGCCA ATGCCCAAGA TAAGCCAATT TTGAGTCTTG GAGGAAAGCC GCTGGTGGGC 3900
TTGGAGGTCA TCAAAAAAAC CACCCATCCC CCTACCACTA CCATGCAGCC CACCACTACT 3960
ACGACGCCCC TGCTTACCAC TACAACCCCG AGGCCACCA CTGCCACCA CATGCGCCCC 4020
ACCACTACTA CGACGCCCCC GCTTACCACT ACACCGAGGC CCACCACTGC CACCAACCCG 4080
CGCACGACCA CAGGCGTCC AACACCCACA GTCCGAACCA CTACGCGGAC AACCAACACC 4140
ACCAACCCCA AACCCACCAC TCCCATCCCC ACCTGTCCCC CTGGACCTTT GGAACCGCAC 4200
GAGATGATG GCAACCTGAT AATGAGCTCC AATGGGATCC CAGAGTGCTA CGCTGAAGAA 4260
GATGAGTCTC CAGGCTTGGG GACTGACACT GCAGTACCTA CGGAAGAGGC CTACGTTATA 4320
TATGATGAAG ATTATGAATT TGAGACGTCA AGGCCACCAA CCACCACTGA GCCTTCGACC 4380
ACTGCTACCA CACCGAGGAT GATCCAGAG GAAGGCGCCA TCAGTCTCTT TCCTGAAGAA 4440
GAATTGATC TGGCTGGAAG GAAACGATT GTTGTCTCTT ACGTGACGTA CCTAAATAAA 4500
GACCCATCAG CCGCGTCTCT TCTGACTGAT GCACCTGGATC ACTTCCAAGT GGACAGCCTG 4560
GATGAAATCA TCCCAATGCA CCTGAAGAAG AGTATCTGCG CTCGCCAGCA TGCTCCCCGC 4620
AACATCACCG TGGTGGCGGT GGAAGGTTGC CACTCATTTG TCATTGTGGA TTGGGACAAA 4680
GCCACCCAG GAGATTGGT CACAGTTAT TTGGTTTACA GTGCATCTTA TGAAGATTTC 4740
ATCAGGAACA AGTTTTCAC TCAAGCTTCA TCAGTAACTC ACTTGCCCAT TGAGAACCTA 4800
AAGCCCAACA CGAGGTATTA TTTTAAAGTG CAAGCAGAAA ATCCTCATGG CTACGGACCT 4860
ATCAGCCCTT CGGTCTCTCT TGTCACCGAA TCAGATAATC CTCTGCTTGT TGTGAGGGCC 4920
CCAGGCGGTG AGCTATCTGG ATCCCATTCG CTTCACAAAC TGATCCAGC TACACGGACT 4980
GCCATGGAGC GCAATATGTG AAGCGCACGT GGTATCGAAA GTTCGTGGGA GTTGTCTTTC 5040
GTAATTCACT GAGGTATAAA ATCTACCTCA GTGACAACT GAAAGATACA TTCTACAGCA 5100
TTGGAGACAG CTGGGAAGA GGTGAAGACC ATTGCCAAT TTGTGATTCA CACCTTGATG 5160
GAAGAACAGG GCCTCAGTCC TATGTAGAAG CCGTCCCTAC TATTCAGGCG TACTATCGCC 5220
AGTATCGTCA GAGCTCTGTC AGGTTTGGGA ACATCGGCTT CGGAACCCCG TACTACTATG 5280
TGGGCTGTA CGAGTGGGG GTCTCCATCC CTGGAAGATG GTAATCACAG GACCGTCATG 5340
CTGCAAGCTT GCCCTGCCCA CCCCCACCAA CTAAGTCGCA CTAGGGGCTG TGAGCAAGA 5400
CAGCCAGCAT GCTCAGCCCC GCTGCCCTAG GTGCCAGGAA GGTACAGAT GGACACTGGC 5460
CATTCGTGTC ATCTCAGTCT GGAACCTCAG CCCACTTCTT GGCGTGACA ATGAACAGGA 5520
TTCACTTTTG TGTGTAATCT TGCTTCTCTA CTCTTTTCTG TTTGTTTGTG ATAGCACATC 5580
CCAGAGACAT CAGAACACAG CAACTGATTC AGTGTGATT CCCAGACTTT TTAGGCATGA 5640
AATTCGAGCA CTTCAGTATT TCCAGGAATA GCATATGCAC GCTGTTCTTG CTTCATGGAA 5700
TGCTACATGC TTTCTGTTT TCTCATTTTG GATTCTTCCA AAACCTAAGT AATTTAAGCT 5760
TCAGGTCCTT TTGTATGAG TAGAAGGAA TATTAAAAA CACCACAAA GAAATAAAT 5820
ATATCTACT TGAATTTTAC TCTATGGACT TACCCACTGC TAGAATAAAT GTATCAAAT 5880
TTATTGTAA ATTCTCAATT TTGATATATA TATGTATATA TGATATACA TATCCACACT 5940
TGCTGTCAA AATATTGATT AAAATTGCTA AATTGTACT TGTTACCAA AAAAAAAA 6000
AAAAAA
  
```

Seq ID NO: 419 Protein sequence
Protein Accession #: Eos sequence

70
 75
 80
 85

```

1 11 21 31 41 51
MPGTLKLRG APADYRVILK TSQEDELDPV DDISVRVMSS QSVLVSNDVP VLEKQKVVVA 60
SRQYTVRYRE KGLARWDYK QIANRRVLIE NLIPDVTYEF AVRIQGERD GKWSTSVFQR 120
TPESAPTTAP ENLNVWVPNG KPTVVAASWD ALPETEGVKV VCLLDLTGLFS VSSFPQSAKS 180
PQNTFFHTPR LSNHLEQSPS PILETLLEPW WMVCSLGNAI FSKSGPQTGE AWDLTPKPSL 240
SLCQCECSTL QKDFSLAYL IDIQTKQVKN DPQLEGSVFG PCFLPYFLTE MLDIGGFSFI 300
MCYEDPVSSL TGNLSKSVAA SKADVQONTE DNGKPEKPEP SSPSPRAPAS SQHPSVPASP 360
QGRNAKDLLE DLKNIKILANG GPRKPQLRA KKAEBELDQS TEITGEBELG SREDSFMSPS 420
DTQDQKRLR PPSRHGHSVV APGRVAVRAR MPALPRRGV DKPQPSLATQ PRFGAPPSAS 480
ASPAHHASTQ GTSRHPSLPA SLNDNDLVDS DEDERAVGSL HPGKGAFAQR PALSPSRQSP 540
SSVLDRSSV HPGAKPASPA RPTPHSGAAE EDSSASAPPS RLSPPHGGSS RLLPTQPHLS 600
SPLSKGGKDG EDAPATNSNA PSRSTMSSSV SHLSSRTQV SEGAASDGE SHGDGDREDG 660
GRQAEATQT LRARPASGHF HLLRHKPFPA NGRSPSRPSI GRGPRLPQSS SPQSTVPSRA 720
HPRVPSHSDS HPKLSGGIHG DEEDEKPLPA TVVNDHVPSS SRQPSISRGWE DLRRSPQRGA 780
SLERKEPIPE NPKSTGADTH PQGKYSSLAS KAQDVQQSTD ADTEGHSFPA QPGSTDHRAH 840
PARPPAARSQ QHPSVPRMT PGRAPEQPP PPVATSQHEP GPQSRDAGRS PSQPRLSLTQ 900
AGRPRPTSQG RSHSSDPYTT ASSRGMLEPTA LQNDQEDAQG SYDDDDSTEVE AQDVRAPAHA 960
  
```

ARAKEAAASL PKHQVBSPT GAGAGGDSR QRGHAASPAR PSRPGGPQR ARVPSRAAPG 1020
 KSEPPSKRPL SSKSQSVSA EDEEEDAGF FRGGKEDLLS SSVPRWPSSS TPRGGKDADG 1080
 SLAKEEREPA LALAPRGSL APVKRPLPPP PGSSPRASHV PSRPPPSRAA TVSPVAGTHP 1140
 WPRYTTRAPP GHFSTTPMLS LRQRMGHARE RNPLSRQPAR PSYRQGYNGR PNVEGKVLPG 1200
 5 SNGKPNQRI INGPQGTWV VDLDRGLVLN AEGRYLQDSH GNPLRIKLG DGRTIVDLGG 1260
 TPVSPDGLP LFGQGRHGT LANAQDKPIL SLGGKPLVGL EVIKKTHHP TTTMQPTTTT 1320
 TPLPTTTTPR PTTATTMQPT TTTTLPITTT PRPTATTTRR TTRRPTTTV RTTTRTTTTT 1380
 TPKPTTPIPT CPFGTLERHD DDGNLIMSSN GIPECYAED EFSGLTDTA VPTEAYVIY 1440
 10 DEDYEFETSR PPTTTEPSTT AITPRVIEPE GAISFPPEEE FDLAGRKRFE APYVTYLNKD 1500
 PSAPCSLTDA LDHFQVDSL EIIIPNDLKS DLPPQHAPRN ITVVAVEGCH SFVIVDWDKA 1560
 TPGDLVTGYL VYSASYEDFI RNKFTQASS VTHLPIENLK PNTRYFVKVQ AQNPFGYGPI 1620
 SPSVSFTVES DNPLLVRRP GGELSGSHSL SNMIPATRIA MDGNM

Seq ID NO: 420 DNA sequence
 Nucleic Acid Accession #: NM_022743
 Coding sequence: 128..1237

1 11 21 31 41 51
 20 GTGGATTTTA GAGATACCTC CCTCCTTCT GCTCAGCTGC CTTCAGTAA TTAACCTCTT 60
 TCTCTGCTGC AACACCCCTA CTGTTCTCG TGTATTGGCT TTTCTGGCA GCAGGAAGGA 120
 AAAGCTGATG CGATGCTCTC AGTGCCGGT CGCCAAATAC TGTAGTGCTA AGTGTCAGAA 180
 AAAAGCTTGG CCAGACCACA AGCGGGAATG CAAATGCCTT AAAAGCTGCA AACCCAGATA 240
 25 TCCTCCAGAC TCCGTTCGAC TTCTTGGCAG AGTTGTCTTC AAACCTATGG ATGAGGACCC 300
 TTCAGATCA GAGAAGCTTT ACTCATTTTA TGATCTGGAG TCAATATTA ACAACTGAC 360
 TGAAGATAAG AAGAGGGGCC TCAGGCAACT CGTAATGACA TTTCAACATT TCATGAGAGA 420
 AGAAATACAG GATGCTCTC AGCTGCCACC TGCCCTTGAC CTTTITGAAG CCTTTGCAAA 480
 AGTGATCTGC AACTCTTTCA CCATCTGTAA TGCGGAGATG CAGGAAGTTG GTGTTGGCCT 540
 ATATCCAGT ATCTCTTTGC TCAATCACAG CTGTGACCCC AACTGTTCGA TTGTGTTCAA 600
 30 TGGGCCCCAC CTCTTACTGC GAGCAGTCCG AGACATCGAG GTGGGAGAGG AGCTCACCAT 660
 CTGCTACCTG GATATGCTGA TGACCACTGA GGAGCGCCGG AAGCAGCTGA GGGACAGTA 720
 CTGCTTTGAA TGTGACTGTT TCCGTTGCCA AACCCAGGAC AAGGATGCTG ATATGCTAAC 780
 TGGTGATGAG CAAGTATGGA AGGAAGTTCA AGAATCCCTG AAAAAAATTG AAGAACTGAA 840
 35 GGCACACTGG AAGTGGGAGC AGGTTCTGGC CATGTGCCAG GCGATCATAA GCAGCAATTC 900
 TGAACGGCTT CCGGATATCA ACATCTACCA GCTGAAGGTG CTCGACTGCG CCATGGATGC 960
 CTGCATCAAC CTCGGCTGTG TGGAGGAAGC CTGTCTCTAT GGTACTCGGA CCATGGAGCC 1020
 ATACAGGATT TTTTCCAGG GAAGCCATCC CGTCAGAGGG GTTCAAGTGA TGAAGTTGG 1080
 CAAACTGCAG CTACATCAAG GCATGTTTCC CCAAGCAATG AAGAATCTGA GACTGGCTTT 1140
 40 TGATATTATG AGAGTGACAC ATGGCAGAGA ACACAGCCTG ATTGAAGATT TGATTCTACT 1200
 TTTAGAAAGA TGCGACGCCA ACATCAGAGC ATCCTAAGGG AACGCAGTCA GAGGGAATA 1260
 CGGCGTGTGT CTTTGTGAA TGCCTTATTG AGGTACACA CTCATGCTT TGTAGCTGT 1320
 GTGAACCTCT CTTATTGAA ATTCTGTTC GTGTTTGTGT AGGTAAATAA AGGCAGACAT 1380
 GOTTITGCAA CCACAGAAT CATTAGTTGT AGAGAAGCAC GATTATAATA AATTCAAAAC 1440
 ATTTGGTTGA GGATGCCAAA AAAAAAATA AAAAAA

Seq ID NO: 421 Protein sequence
 Protein Accession #: NP_073580

1 11 21 31 41 51
 50 MRCSQCRVAK YCSAKQKKA WPDHKRECK LKSCPKPRYP DSVRLLEGRV FKLMDGAPSE 60
 SEKLYSPYDL ESNINKLTD KKEGLRLQVM TFQHFMRREE QDASQLPPAP DLFEAFARVI 120
 CNSFTICNAE MQEVGVGLYP SISLLNHSCD PNCISVFNPG HLLLRVRDI EVGEELTICY 180
 55 LDMLMTSEER RQRLRDQYCF ECDCEFCQTO DKDADMLTGD EQVMKEVQES LKKIEELKAH 240
 WKWEQVLAMC QAISSNSER LPDINIYQLK VLDCAMDACI NLGLLEALF YGTRTMEPYR 300
 IFPPGSHFVR GVQVMKVGLK QLHQGMFPQA MNRLRLAFDI MRVTHGREHS LLEDLILILE 360
 ECDANIRAS

Seq ID NO: 422 DNA sequence
 Nucleic Acid Accession #: NM_003014.2
 Coding sequence: 238..648

1 11 21 31 41 51
 65 GCGGGTTTGG CGCCCGAAG GCTGAGAGCT GCGGCTGCTC GTGCCCTGTG TGCCAGACGG 60
 CGGAGCTCCG CGGCGGACCC CCGCGGCCCC GCTTTGCTGC CGACTGGAGT TTGGGGGAAG 120
 AAATCTCTCT CGGCGCCAGA AGATTCTTTC CTGCGCGAAG GGACAGCGAA AGATGAGGGT 180
 GGCAGGAAGA GAGGCGCTT TCTGTCTGCC GGGGTGCGAG CGCGAGAGGG CAGTGCCATG 240
 70 TTCTCTCCA TCCTAGTGGC GCTGTGCTGT TGGCTGCACC TGGCGCTGGG CGTGCAGGGC 300
 GCGCCCTGGG AGGCGGTGGC CATCCCTATG TGCCGCGACA TGCCCTGGAA CATCACGGG 360
 ATGCCCAACC ACCCTGACCA CAGCACGCGAG GAGAAGCGCA TCCTGGCCAT CGAGCAGTAC 420
 GAGGAGCTGG TGGAGCTGAA CTGCAGCGCC GTGCTGCGCT TCTTCTCTG TGCCATGTAC 480
 GCGGCCATTG GCACCTGGA GTTCTGTCAC GACCTATCA AGCCGTGCAA GTCGGTGTGC 540
 75 CAACGCGCGC GOGAGAGTGC CGAGCCCTCT ATGAAGATGT ACAACCCAG CTGCGCCGAA 600
 AGCCTGGCGC GCGAGAGTCT GCCTGTCTAT GACCGTGGGG TGTGCATTTC GCCTGAAGCC 660
 ATCGTCAAGG ACCCTCCGGA GGATGTTAAG TGGATAGACA TCACACCAGA CATGATGGTA 720
 CAGGAAGAGC CTCTTGATGT TGACTGTAAA CGCCTAAGCC CGATGCGGTG CAAGTGTAAA 780
 AAGGTGAAGC CAACTTTGGC AACGTATCTC AGCAAAACT ACAGCTATGT TATTCTATGC 840
 80 AAAATAAAG GTGTGCAGAG GAGTGGCTGC AATGAGSTCA CAACGGTGGT GGATGTAAAA 900
 GAGATCTTCA AGTCTCATC ACCCATCCCT CGAATCTAAG TCCGCTCAT TACAAATTCT 960
 TCTTCCAGT GTCCACACAT CCTGCCCAT CAAGATGTT TCATCATGTG TTACGAGTGG 1020
 CGTTCAAGGA TGATGCTTCT TGAATTTGCG TTAGTTGAAA AATGGAGAGA TCAGCTTAGT 1080
 AAAAGATCCA TACAGTGGGA AGAGAGGCTG CAGGAACAGC GGAGAACAGT TCAGACAAAG 1140
 85 AAGAAAACAG CCGGGCGCAC CAGTGTAGT AATCCCCCA AACCAAGGG AAAGCTCTCT 1200
 GCTCCCAAC CAGCCAGTCC CAAGAAGAAC ATTAATAACTA GGAGTGCCCA GAAGAGAACA 1260
 AACCCGAAA GAGTGTGAGC TAACTAGTT CCAAGCGGA GACTTCCGAC TTCTTACAG 1320
 GATGAGGCTG GGCATTGCCT GGGACAGCCT ATGTAAGGCC ATGTGCCCCT TGCCCTAACA 1380

	ACTCACTGCA	GTGCTCTTCA	TAGACACATC	TTGTCAGCAT	TTTCTTAAGG	CTATGCTTCA	1440
	GTTTTTCTTT	GTAAGCCATC	ACAAGCCATA	GTGCTAGGTT	TGCCCTTTGG	TACAGAAGGT	1500
	GAGTTAAAGC	TGGTGAAGAA	GGCTTATTGC	ATTGTCATTG	GAGTAACCTG	TGTGCATACT	1560
5	CTAGAAGAGT	AGGGAAATAA	ATGCTTGTTA	CAATTCGACC	TAATATGTGC	ATTGTAAAT	1620
	AAATGCCATA	TTTCAACAA	AACACGTAAT	TTTTTTACAG	TATGTTTAT	TACCTTTTGA	1680
	TATCTGTGT	TGCAATGTGA	GTGATGTTTT	AAATGTGAT	GAAATATATA	TGTTTTTAAG	1740
	AAGGAACAGT	AGTGAATGA	ATGTTAAAG	ATCTTTATGT	GTTTATGGTC	TGCAGAAGGA	1800
	TTTTTGTAT	GAAAGGGAT	TTTTTGAAAA	ATTAGAGAAG	TAGCATATGG	AAAATTATAA	1860
10	TGTGTTTTTT	TACCAATGAC	TTCACTTTCT	GTCTTTAGCT	AGAACTTAA	AAACAAAAAT	1920
	AATAATAAAG	AAAAATAAAT	AAAAAGGAGA	GGCAGACAAT	GTCTGGATTC	CTGTTTTTTG	1980
	GTTACCTGAT	TCCATGATC	ATGATGCTTC	TTGTCACAC	CCTCTTAAGC	AGCACCAGAA	2040
	ACAGTGAGTT	TGCTGTGACC	ATTAGGAGTT	AGGTACTAAT	TAGTTGGCTA	ATGCTCAAGT	2100
	ATTTATACC	CACAAGAGAG	GTATGTCAC	CATCTTACT	CCCAGGACAT	CCACCCTGAG	2160
15	AATAATTGA	CAAGCTTAA	AATGGCCTTC	ATGTGAGTGC	CAAAATTTGT	TTTTCTTCAT	2220
	TTAAATATT	TCTTGCCTA	AATACATGTG	AGAGGAGTTA	AATATAAATG	TACAGAGAGG	2280
	AAAGTTGAGT	TCCACCTCTG	AAATGAGAAT	TACTTGACAG	TTGGGATACT	TTAATCAGAA	2340
	AAAAAGAACT	TATTTGCAGC	ATTTTATCAA	CAAAATTCAT	AATGTGAGAC	AATTTGGAGGC	2400
	ATTTATTTTA	AAAAACAATT	TTATTGGCCT	TTTGCTAAAC	CAGTAAGCAT	GTATTTTATA	2460
20	AGGCATTCAA	TAAATGCACA	ACGCCCAAG	GAAATAAAT	CCTATCTAAT	CCTACTCTCC	2520
	ACTACACAGA	GGTAATCACT	ATTAGTATT	TGGCATATTA	TTCTCCAGGT	GTGCTTAT	2580
	GCACCTATAA	AATGATTGGA	ACAAATAAAA	CTAGGAACCT	GTATACATGT	GTTCATAAC	2640
	CTGCCCTCTT	TGCTTGGCCC	TTTATTGAGA	TAAGTTTTCC	TGTCAGAGAA	GCAGAAACCA	2700
	TCTCATTCT	AACAGCTGTG	TTATATTCCA	TAGTATGCAT	TACTCAACAA	ACTGTTGTGC	2760
25	TATTGGATAC	TTAGGTGGTT	TCTTCACTGA	CAATACTGAA	TAAACATCTC	ACCGGAATTC	

Seq ID NO: 423 Protein sequence
Protein Accession #: NP_003005.1

30	1	11	21	31	41	51	
	MFLSILVALC	LWLHLALGVR	GAPCEAVRIP	MCRHMPNNIT	RMPNHLHST	QENAILAIBQ	60
	YEELVDVNC	AVLRFFFCPE	YAPICTLEFL	HDPKPKCKSV	QQRARDCEP	LMKMYNHSNP	120
	ESLACDELTV	YDRGVCISPE	AIIVDLPELV	KWIDITPDM	VQERPLDVDC	KRLSPDRCKC	180
35	KKVKPLATY	LSKNYSYVH	AKIKAVQRS	CNEVTVVDV	KEIFKSSSPI	PRTVPLITN	240
	SSQCCHILP	HQDVLIMCYE	WRSRMMLLEN	CLVEKWRDQL	SKRSIQWEER	LQEQRRVTQD	300
	KKKTAGRTSR	SNPPKPKGKP	PAPKASPCK	NIKTRSAQKR	TNPKRV		

Seq ID NO: 424 DNA sequence
Nucleic Acid Accession #: BC010423
Coding sequence: 248..1780

40	1	11	21	31	41	51	
45	CACAGCGTGG	GAAGCAGCTC	TGGGGGAGCT	CGGAGCTCCC	GATCAGGCT	TCTTGGGGGT	60
	AGCTACGGCT	GGGTGTGTAG	AACGGGGCCG	GGGCTGGGGC	TGGTCCCTT	AGTGGAGACC	120
	CAAGTGCGAG	AGGCAAGAAC	TCTGCAGCTT	CCTGCCCTCT	GGGTCACTTC	CTTATTCAAG	180
	TCTGCAGCCG	GCTCCGAGGG	AGATCTCGST	GGAACCTCAG	AAACGCTGGG	CAGTCTGCGT	240
	TTCAACCATG	CCCTGTCTCC	TGGGAGCCGA	GATGTGGGGG	CCTGAGGCTT	GGCTGTGCT	300
50	GCTGTACTGT	CTGGCATCAT	TTACAGGCCG	GTGCCCCCGG	GGTGAGCTGG	AGACCTCAGA	360
	CGTGTAACT	GCGTGCATGT	GCCAGGACGC	AAAACCTGCC	TGCTTCTACC	GAGGGGATCT	420
	CGGCGAGCAA	GTGGGCAAG	TGGCATGGGC	TCCGGTGGAC	GCGGCGGAAG	GCGCCAGGAA	480
	ACTAGCGCTA	CTGCACTCCA	AATACGGGCT	TCAATGTGAG	CCGCTTACG	AGGGCCGCGT	540
	GGAGCAGCCG	CGGCCCCAC	GCAACCCCT	GGACGGCTCA	GTGCTCCTGC	GCAACGCACT	600
55	GCAGGCGGAT	GAGGGCGAGT	ACGAGTGCCG	GGTCAGCACC	TTCCCGCCCG	GCAGCTTCCA	660
	GGCGGGCTG	CGGCTCCGAG	TGCTGGTGCC	TCCCTGCGCC	TCACTGAATC	CTGGTCCAGC	720
	ACTAGAAGAG	GGCCAGGGCC	TGACCTGGG	AGCCTCCTGC	ACAGCTGAGG	GCAGCCGAGC	780
	CCOCAGCGTG	ACCTGGGACA	CGGAGGTCAA	AGGCACAAG	TCCAGCCGTT	CCTTCAAGCA	840
	CTCCCGCTCT	GCTGCCGTCA	CCTCAGAGTT	CCAATTGGTG	CCTAGCCGCA	GCATGAATGG	900
60	GCAGCCACTG	ACTTGTGTGG	TGTCCCATCC	TGGCCTGCTC	CAGGACCAAA	GGATCAACCA	960
	CATCTCCAC	GTGTCTCTCC	TTGCTGAGGC	CTCTGTGAGG	GGCCTTGAAG	ACCAAAATCT	1020
	GTGGCACAAT	GGCAGAGAAG	GAGCTATGCT	CAAGTGCCCT	AGTGAAGGGC	AGCCCCCTCC	1080
	CTCATACAAC	TGGACACGCG	TGGATGGGCC	TCTGCCAGT	GGGTACGAG	TGGATGGGGA	1140
	CACCTTTGGGC	TTTCCCCCAC	TGACCACTGA	GCACAGCGCG	ATCTACGTCT	GCCATGTGAG	1200
65	CAATGAGTTC	TCCTCAAGGG	ATTCTCAGGT	CACGTGGGAT	GTCTTGACC	CCCAGGAAGA	1260
	CTCTGGGAAG	CAGGTGGACC	TAGTGTACG	CTCGGTGGTG	TGGTGGGGTG	TGATCGCCGC	1320
	ACTCTGTGTC	TGCTTCTGCG	TGGTGGTGGT	GGTGTCTCAT	TCCCGATACC	ATCGGCGCAA	1380
	GGCCCGCAGC	ATGACCCAGA	AATATGAGGA	GGAGCTGACC	CTGACCAAGG	AGAACTCCAT	1440
	CCGGAGGCTG	CATTCCCATC	ACACGGACCC	CAGGAGCCAG	CCGAGGAGA	GTGTAGGGCT	1500
70	GAGAGCCGAG	GGCCACCCCT	ATAGTCTCAA	GGACAACAGT	AGCTGTCTGT	TGATGAGTGA	1560
	AGAGCCCGAG	GGCCGCGAGT	ACTCCACGCT	GACCAAGGTT	AGGGAGATAG	AAACACAGAG	1620
	TGAATCTGCT	TCTCCAGGCT	CTGGGCGGGC	CGAGGAGGAG	GAAGATCAGG	ATGAAGGCAT	1680
	CAAAACAGGC	ATGAACCAAT	TTGTTCAGGA	GAATGGGACC	CTACGGGCCA	AGCCACGGG	1740
	CAATGGCATC	TACATCAATG	GGCGGGGACA	CCTGGTCTGA	CCCAGGCTGT	CCTCCCTTCC	1800
75	CTAGGCGCTG	CTCTCTCTGT	TGACATGGGA	GATTTTAGCT	CATCTTGGGG	GCTCTCTTAA	1860
	ACACCCCATC	TTCTTGGCGA	AGATGCTCCC	CATCCCACTG	ACTGCTTGAC	CTTAACTCTC	1920
	AACCTTCTGT	TTATCGGGA	GGGCTCCACC	AATGAGTCT	CTCCCACTAT	GCATGCAAGT	1980
	CACGTGTGT	GTGATGTGT	GCCTGTGTGA	GTGTTGACTG	ACTGTGTGTG	TGTGGAGGGG	2040
	TGACTGTCCG	TGGAGGGGTG	ACTGTGTCCG	TGGTGTGTAT	TATGCTGTCA	TATCAGAGTC	2100
80	AAGTGAACCT	TGTTGTATGT	GCCACGGGAT	TTGAGTGGTT	GCGTGGGCAA	CACGTGTCAG	2160
	GTTTGGCGTG	TGTGTCTGT	GGCTGTGTGT	GACCTCTGCC	TGAAAAGACA	GATATTCTCT	2220
	CAGACCCAG	AGCAGTATTA	ATGATGCAGA	GGTTGGAGGA	GAGAGGTGGA	GACTGTGGCT	2280
	CAGACCCAGG	TGTGCGGGCA	TAGCTGGAGC	TGGAATCTGC	CTCCGTGTGT	AGGGAACCTG	2340
	TCTCCTACCA	CTTCGGAGCC	ATGGGGGCAA	GTGTGAAGCA	GCCAGTCCCT	GGGTGAGCCA	2400
	GAGGCTTGAA	CTGTTACAGA	AGCCCTCTGC	CCTCTGGTGG	CCTCTGGGCG	TGCTGCAATG	2460
85	ACATATTTTC	TGTAAATATA	CATGCGCGGG	GAGCTTCTTG	CAGGAATACT	GCTCCGAATC	2520
	ACTTTTAAAT	TTTTTCTTTT	TTTTTCTTGT	CCCTTTCCAT	TAGTGTATTT	TTTTATTAT	2580
	TTTTATTATT	ATTTTATTTT	AGAGTTTGAG	TCCAGCCTGG	ACGATATAGC	CAGACCCCTG	2640

CTGTAAAAAA ACCAAAAACC AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 425 Protein sequence

Protein Accession #: AAH10423

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

1 11 21 31 41 51
MPLSLGAEMW GPEAWLLLLL LLASFTGRCP AGELETSDDV TVVLGQDAKL PCFYRGDSGE 60
VQGVQAWARV DAGGAQELA LHSKYGLHV SPAYEGRVEQ PPPRNPLDG SVLLRNAVQA 120
DEGSYECRVS TFPAGSFQAR LRLRVLPPL PSLNPGPALE EQQLTLAAS CTAECSFAPS 180
VTWDTVEKGT TSSRSFKHSR SAAVTSEFHL VPSRSMNGQP LTCVVSHPGL LQDQRITHIL 240
HVSFLAEASV RGLDQNLWH IREGAMKLC LSEGQPPPSY NWTRLDGLPL SGVRVDGDTL 300
GFPLPTEHS GIYVCHVSNE FSSRDSQVTV DVLDPOEDSG KQVDLVASV VVGVIAALL 360
FCLLVVVVVL MSRYHRRKAQ QMTQKYESEL TLTRENSIRR LHSHTDPRS QPEESVGLRA 420
EGHPDSLKDN SSCSVMSEEP EGRSYSTLT VREIETQTEL LSPGSGRAEB EEDQDEGIKQ 480
AMNHFVQENG TLRAKPTGNG IYINGRHLV

Seq ID NO: 426 DNA sequence

Nucleic Acid Accession #: NM_003474.2

Coding sequence: 37..3036

1 11 21 31 41 51
CACTAACGCT CTCTCTAGTC CCCGGGCCAA CTCGGACAGT TTGCTCATT ATTGCAACGG 60
TCAAGGCTGG CTGTGTCCAG AACGGCGCGC GCGCGACGCA CGCACACACA CGGGGGGAAA 120
CTTTTTTAA AATGAAGGC TAGAAGAGCT CAGCGCGCGC GCGGGCCGTG CGCGAGGGCT 180
CCGGAGCTCA CTCGCGAGG CAGGAAATCC CTCGGTCCG GACGCCCGGC CCCGCTCGGC 240
GCCCGGTGG GATGTCGAG CGCTCGCGC GGGGCCCGAG AGCTGCTGCA CTGAAGGCGC 300
GCGACGATGG CAGCGCGCC GCTGCCCGTG TCCCCCGCC GCGCCCTCT GCTCGCCCTG 360
GCCGGTGTCT TGCTCGCGCC CTGCGAGGCC CGAGGGGTGA GCTTATGGAA CGAAGGAAGA 420
GCTGATGAAG TTGTCTAGTC CTCTGTTCGG AGTGGGGACC TCTGGATCCC AGTGAAGAGC 480
TTGAGCTCCA AGAATCATCG AGAAGTGTCT AATATTGAC TACAACGGGA AAGCAAGAA 540
CTGATCATAA ATCTGGAAG AAATGAAGGT CTCATTGCCA GCAGTTTTC GGAACCCAC 600
TATCTGCAAG ACGGTACTGA TGCTCCCTC GCTCGAAAT ACACGGTAAT TCTGGGTCTC 660
TGTTACTACC ATGGACATGT ACGGGGATAT TCTGATTGAG CAGTCAGTCT CAGCAGCTGT 720
TCTGGTCTCA GGGGACTTAT TGTTTGA AATGAAAGCT ATGCTTTAGA ACCAATGAAA 780
AGTGCAACCA ACGATACAA ACTCTTCCA GCGAAGAAGC TGAAAAGCGT CCGGGGATCA 840
TGTTGATCAC ATCAACAC ACCAACCTC GCTGCAAGA ATGTTTTC ACCACCTCT 900
CAGACATGG CAAGAAGCA TAAAGAGAG ACCCTCAAG CAACTAAGTA TGTGGAGCTG 960
GTGATGTTGG CAGACAACCG AGAGTTTCAG AGGCAAGGAA AAGATCTGGA AAAAGTTAAG 1020
CAGCGATTAA TAGAGATTGC TAATCAGCTT GACAAGTTT ACAGACCATT GAACATTCGG 1080
ATGTTGTTGG TAGGCTGGA AGTGTGGAAT GACATGGACA AATGCTCTGT AAGTCAGGAC 1140
CCATTACCA GCCTCCATGA ATTCTGGAG TGGAGGAAGA TGAAGCTTCT ACCTCGCAAA 1200
TCCCATGACA ATGCGCAGCT TGTCAGTGGG GTTTATTTC AAGGACCAC CATCGCATG 1260
GCCCAATCA TAGCATGTG CACGGCAGAC CAGTCTGGGG GAATTGTGAT GACCATTTCA 1320
GACAATCCCC TTGTCGACG CGTGACCTGT GCACATGAGC TGGGCCACAA TTTCGGGATG 1380
AATCATGACA CACTGACAG GGGCTGTAGC TGTCAAATGG CGTTTGAGAA AGGAGGCTGC 1440
ATCATGAACG CTTCCACCG GTACCCATT CCCATGGTGT TCAGCAGTTG CAGCAGGAAG 1500
GACTTGGAGA CCAGCCTGGA GAAAGGAAT GGGGTGTGCC TGTTTAACT GCCGGAAGTC 1560
AGGGAGTCTT TCGGGGCCA GAAGTGTGGG AACAGATTG TGGAAGAAG AGAGAGTGT 1620
GACTGTGGGG AGCCAGAGAA ATGTATGAAT CGCTGTGCA ATGCCACCA CTGTACCTGT 1680
AAGCCGAGCG CTGTGTGCGC ACATGGGCTG TGCTGTGAAG ACTGCCAGCT GAAGCCTGCA 1740
GGAACAGCGT GCAGGACTCC CAGCACTCC TGTCAGCTCC CAGAGTCTG CACAGGGGCC 1800
AGCCCTCACT GCCAGCCAA CGTGTACTG CACAGTGGGC ACTCATGTCA GGATGTGGAC 1860
GGCTACTGCT ACAATGGCAT CTGCCAGACT CACGAGCAGC AGTGTGTCTC ACTCTGGGA 1920
CCAGGTGCTA AACCTGCCCC TGGGATCTGC TTTGAGAGAG TCAATTCTGC AGTGATCTCT 1980
TATGGCAACT TGGCAAGAG CTCGAAGAGT TCCTTTGCCA AATGCGAGAT GAGAGATGCT 2040
AAATGTGAA AATCAGCATG TCAAGGAGGT GCCAGCCGCG CAGTCATTGG TACCAATGCC 2100
GTTTCCATAG AAACAAACAT CCCCCTGCAG CAAGGAGGCC GGATTCTGT CCGGGGAC 2160
CACGTGTACT TGGGCGATGA CATGCCGGAC CCAGGGCTTG TGCTTGCAGG CACAAGTGT 2220
GCAGATGAAA AATCTGCTCT GAATCTGCAA TGTCAAAATA TTAGTGTCTT TGGGTTTCA 2280
GAGTGTGAAA TGCAGTGCCA CGGCAGAGGG GTGTGCAACA ACAGGAAGAA CTCGCACTGC 2340
GAGGCCACT GGGCCTCTCC CTTCTGTGAC AAGTTTGGCT TTGGAGGAG CACAGACAGC 2400
GGCCCCATCC GGCAGCAGA TAAACAAAGT TTAACCATAG GAATTCTGGT GACCATCTG 2460
TGCTTCTTG CTGCCGATT TGTTGTTTAT CTCAAAAGGA AGACCTTGAT ACAGTCTGT 2520
TTTCAAAATA AGAAGACCAC CATTGAAAAA CTAAGGTGTG TGGGCCCTTC CGGCCACCC 2580
CGTGGCTTCC AACCTGTCA GGCTCACCTC GGCCACCTTG GAAAAGGCTT GATGAGGAAG 2640
CGGCCAGATT CCTACCCACC GAAGGACAA: CCCAGGAGAT TGCTGCAGTG TCAGAATGTT 2700
GACATCAGCA GACCCCTCAA CGGCCTGAAT GTCCTCAGC CCCAGTCAAC TCAGCGAGTG 2760
CTTCTCTCCC TCCACCGGGC CCCACGTGCA CCTAGCGTCC CTGCCAGACC CCGCCAGCC 2820
AAGCCTGCAC TTAGGCAGGC CCAGGGGACC TGTAGGCCAA ACCCCCTCA GAAGCCTCTG 2880
CCTGCAGATC CTCTGGCCAG AACAACCTCG CTCACTCATG CCTTGGCCAG GACCCAGGA 2940
CAATGGGAGA CTGGGCTCCG CTGGCACCCT CTCAGACCTG CTCACAATA TCCACACCA 3000
GTGCCAGAT CCACCCACAC CGCCTATATT AAGTGAGAAG CCGACACCTT TTTTCAACAG 3060
TGAAGACAGA AGTTTGCACT ATCTTTCAGC TCCAGTTGGA GTTTTTTGA CCAACTTTTA 3120
GGATTTTTTT TAATGTTTAA AACATCATT CTATAAGAAC TTTGAGTAC TGCCGTCACT 3180
GCTGTGCTGT GCTATGGTGC TCTGTCTACT TGCAACAGTA CTTGTAATTT ATTAATTTAT 3240
GCAGATGTT GATTACAGTG CAGTGCCTG TAGTAGGCAT TTTTACCAT ACTGAGTTT 3300
CCATGGCAGG AAGCTTGTG TTGCTTTTAG TATTTAGTG AACTTGAAT ATCCTGCTTG 3360
ATGGGATTCT GGACAGATG TGTTGCTTT CTGATCAAGG CCTTATTGGA AAGCAGTCCC 3420
CCAACTACCC CCAAGCTGTG TTAATGTACC AGATGCACT CAAGAGATCC CAAGTAGAAT 3480
CTCAGTTGAT TTTCTGGATT CCCCATCTCA GGCCAGAGCC AAGGGGCTTC AGGTCCAGGC 3540
TGTGTTTGGC TTTCAGGAG GCCCTGTGCC CCTTGACAAC TGGCAGGAG GCTCCAGGG 3600
ACACCTGGGA GAAATCTGCG TTCTGGCCAG GAAGCTTTGG TGAGAACCTG GGTTCAGAC 3660
AGGAATCTTA AGGTGTAGCC ACACAGGAT AGAGACTGGA ACACAGACA AGCCAGAAT 3720
TGACCCCTGAG CTGACCAACC GTGAGCATGT TTGGAAGGGG TCTGTAGTGT CACTCAAGGC 3780
GGTGTCTGAT AGAATGCCA AGCACTTCTT TTTCTGCTG TCCTTTCTAG AGCACTGCCA 3840

5 CCAGTAGGTT ATTTAGCTTG GGAAAGGTGG TGTTCCTGTA AGAAACCTAC TGCCCAAGCA 3900
 CTGCAAAACCG CCACCTCCCT ATACTGCTTG GAGCTGAGCA AATCACCACA AACTGTATAA 3960
 CAATGATCCT GTATTGAGAC AGATGAGGAC TTTCCATGGG ACCACAACCTA TTTTCAGATG 4020
 TGAACCATTA ACCAGATCTA GTCAATCAAG TCTGTTTACT GCAAGGTTCA ACTTATTAACT 4080
 AATTAGGCGAG ACTCTTTATG CTTGCAAAAA CTACAACCA TGGAAATGTA GTTTCATGGG 4140
 TATAGTTTAT GTCTGCTATC ATTATTGCTA GATATTGGAC AAAGAACCCTT CTCTATGGGG 4200
 CATCCTCTTT TTCCAACCTG GCTGCAGGAA TCTTTAAAG ATGCTTTTAA CAGAGTCTGA 4260
 10 ACCTATTCTT TAAACACTTG CAACCTACCT GTTGAGCATC ACAGAATGTG ATAAGGAAAT 4320
 CAACTTGCTT ATCAACTTCC TAAATATTAT GAGATGTGGC TTGGGCAGCA TCCCTTGAA 4380
 CTCTTCACTC TTCAAATGCC TGACTAGGGA GCCATGTTTC ACAAGGTCTT TAAAGTGA 4440
 AATGGCATGA GAAATACAAV AATACTCAGA TAAGGTAAAA TGCCATGATG CCTCTGTCTT 4500
 CTGGACTGTT TTTCACTATA GAAGACAATT GACACAGTT ACATAATTCA CTCTGAGTGT 4560
 TTTATGAGAA AGCCTTCTTT TGGGGTCAAC AGTITTCCTA TGCTTTGAAA CAGAAAAATA 4620
 TGTACCAAGA ATCTTGGTTT GCCTTCCAGA AAACAAAAC TGCATTCACT TTCCCGGTGT 4680
 15 TCCCACTGTT ATCTAGGCAA CATAGTATTC ATGACTATGG ATAACTAAA CACGTGACAC 4740
 AAACACACAC AAAAGGGAAC CCAGCTCTAA TACATTCCAA CTGATATAGC ATGCATCTGT 4800
 TTATTCTATA GTTATTAAAG TCTTTAAAT GTAAAGCCAT GCTGGAAAAT AATACTGCTG 4860
 AGATACATAC AGAATTACTG TAACTGATTA CACTTGTAAT TTGTACTAAA GCCAAACATA 4920
 20 TATATACTAT TAAAGAGTT TACAGATTT TATGTTGCAT TACGTGGGCA TTGCTTTTT 4980
 AGATGCCCAA ATCTCTAGAT CTGGCATGTT AGCCCTTCCT CCAATTATAA GAGGATATGA 5040
 ACCAAAAAAA AAAAAAAA AA

Seq ID NO: 427 Protein sequence
Protein Accession #: NP_003465

25 1 11 21 31 41 51
 MAARPLVPSP ARALLLALAG ALLAPCEARG VSLWNEGRAD EVVSASVRSQ DLWIPVKSFD 60
 SKNHPEVLM RLORESKELI INLERNEGLI ASSFTETHYL QDGTDVSLAR NYTVILGHY 120
 30 YHGHVRGYS SAVSLTSCSG LRGLIVFENE SYVLEPMKSA TNRYKLFPAK KLSVRSQSG 180
 SHHTPNLAA KNVFPFSPQT WARRHKRETL KATKYVELVI VADNREFQRO GKDLKVKOR 240
 LIBIANHVOK FYRFLNIRIV LVGVEVNDM DKCSVSQDPF TSLHEFLDWR KMKLLPRKSH 300
 DNaQLVSGVY FQGLTIGMAP IMSMCTADQS GGIIVMDHSDN PLGAAVTLAH ELGHNFGMNH 360
 35 DTLDRGCSQ MAVKKGCGIM NASTGYPPFM VFSSCSRDL ETSLEKMGV CLFNLPVRE 420
 SPFGQKQGNR FVEGEGECDC GEPEECMNR CNAATCTLKP DAVCAHGLCC EDCQLKPA 480
 ACRDSSNSCD LPEPCTGASP HCPANVYLHD GHSCQDVG DYNGICQTHE QCVTLWPG 540
 AKPAPGICFE RVNSAGDPYG MCGKVSXSSF AKCEMRDACK GKICQCGGAS RPIVGNV 600
 40 IBTNILPQGG GRLLCRGTHV YLGDDMPDPG LVLASTKCAD GKICLNRCQ NISVFGVHEC 660
 AMQCHGRGVC NNRKNCHEA HWAPPFCDKF GFGGSTDGSG IRQADNQLT IGLVLTILCL 720
 LAAGFVVYVK RKTILRLFT NKKTIEKLR CVRPSRPPRG FQPCQHLGH LGKGLMRKPP 780
 DSYPPKDNPR RLQCCQNDVI SRPLNGLNVP PQPQSTQVLP PLHRAFRAPS VPARELPAPK 840
 ALRQAGTCK PNPQKPLPA DPLARTTRLT HALARTPGQW ETGLRLAPLR PAPQYPHQVP 900
 RSTHTAYIK

Seq ID NO: 428 DNA sequence
Nucleic Acid Accession #: NM_003714
Coding sequence: 135..1043

50 1 11 21 31 41 51
 GAGGAGGAGG GAAAAGGCGA GCAAAAAGGA AGAGTGGGAG GAGGAGGGGA AGCGGCGAAG 60
 GAGGAAGAGG AGGAGGAGGA AGAGGGGAGC ACAAGGATC CAGGTCTCCC GACGGGAGGT 120
 TAATACCAAG AACCATGTGT GCCGAGCGGC TGGGCCAGTT CATGACCTCG GCTTTGGTGT 180
 TGGCCACCTT TGACCCGGCG CGGGGAGCCG AGGCCACCAA CCCACCCGAG GGTCCCAAG 240
 55 ACAGGAGCTC CCAGCAGAAA GCGCGCTGT CCTGCAGAA TACAGCGGAG ATCCAGCACT 300
 GTTTGGTCAA CGCTGGCGAT GTGGGGTGTG GCGTGTGTA ATGTTTCGAG AACAACTCTT 360
 GTGAGATTGG GGGCTTACAT GGGATTGCA TGACTTTCT GCACAACCT GGAATAATTG 420
 ATGCCAGGG CAGTCTATT ATCAAAGACG CTTGAAATG TAAGGCCAC GCTCTGCGGC 480
 60 ACAGGTTGGG CTGCATAAGC CGGAAGTGC CGGCCATCAG GGAATGGTG TCCAGTTGC 540
 AGCGGGAATG CTACCTCAAG CACGACCTGT GCGCGGCTGC CCAGGAGAAC ACCCGGGTGA 600
 TAGTGGAGAT GATCCATTTC AAGGACTTGC TGCTGCACGA ACCCTAGCTG GACCTCGTGA 660
 ACTTGCTGCT GACCTGTGGG GAGGAGGTGA AGGAGGCCAT CACCCACAGC GTGCAAGTTC 720
 AGTGTGAGCA GAAGTGGGGA AGCCTGTGCT CCATCTTGAG CTTCTGCACC TCGGCCATCC 780
 65 AGAAGCCTCC CACGCGCCCC CCGGAGCGCC AGCCCCAGGT GGACAGAAC AAGCTCTCCA 840
 GGGCCACCA CGGGGAAGCA GGACATCACC TCCAGAGCC CAGCAGTAGG GAGACTGGCC 900
 GAGGTGCCAA GGTGAGCGGA GGTAGCAAGA GCCACCAA CGCCCATGCC CGAGGCAGAG 960
 TCGGGGCTCT TGGGGCTCAG GGACCTTCOG GAAGCAGCGA GTGGGAAGAC GAACAGTCTG 1020
 AGTATTCTGA TATCCGAGG TGAAATGAAA GGCCTGGCCA CGAAATCTTT CCTCCACGCC 1080
 70 GTCCATTTC TTATCTATGG ACATTCCAAA ACATTACCA TTAGAGAGGG GGGATGTAC 1140
 ACGCAGGATT CTGTGGGACT TGTGGACTTC ATCGAGGTGT GTGTTCCGCG AACCGACAGG 1200
 TGAGATGGAG ACCCTTGGG COGTGGGCTC TCAGGGGTGC CTGGTGAATT CTGCACTTAC 1260
 AGCTACTCAA GGGAGCGCGC COGCTTATC CTCGTACCTT TGTCTCTTT CCATCTGTGG 1320
 AGTCAGTGGG TGTGCGCGCG TCTGTTGTGG GGGAGGTGAA CCAGGGAGGG GCAGGGCAAG 1380
 75 GCAGGGCCCC CAGAGCTGGG CCACACAGTG GGTGCTGGGC CTGCCCCGGA AGCTCTGCTG 1440
 GCAGCAGCCT CTGGTGTCTG CTCGCGGAA GTCAGGGCGG CTGATTCCA GCACAGGAGT 1500
 GAATGTAAAA ATAAATATCG CTTAGAAATG AGGAGAAGGG TGGAGAGGAG GCAGGGGCGG 1560
 AGCGGGTCTT TGGTGCCAAA CTGAAATTCA GTTCTTGTG TGGGGCCTTG CGGTTCAGAG 1620
 CTCTTGGCGA GGGTGGAGGG AGGAGTGTCA TTTCTATGT TAATTTCTGA GCCATTGTAC 1680
 80 TGTCTGGGCT GGGGGGACCA CTGTCCAAGG GAGTGGCCCC TATGAGTTTA TATTTTAAAC 1740
 ACTGCTTCAA ATCTCGATTT CACTTTTTTT ATTATCCAG TGGGTCAATA AAACAGCTC AAAGGGGTT 1800
 TCTAAATAAA TGGCTTCAA ACAAGCAAC TGGGTCAATA AAACAGCTC AAAGGGGTT 1860
 TAAATAAA AAACAGGCC CATCTTTGA GGCTGATTT TCTTTTTTT AAGTCTTATT 1920
 TAAAGCTA TCAACAGCG ACATAGCCAT ACATCTGACT GCCTGACATG GACTCTGCC 1980
 CACTTGGGGG AAACCTTATA CCCAGAGGAA AATACACACC TGGGAGTAC ATTTGACAAA 2040
 85 TTTCCTTAG GATTTCGTTA TCTCACCTTG ACCCTAGCC AAGATTGGTA AAGCTGCTC 2100
 TTGGCGATTC CAGGAGACC CTGGCTTCTC CATGTGAGGG GATGGGAAAG 2160
 GAAAGAGAG AATGAAGACT ACTTAGTAAT TOCCATCAGG AAATGCTGAC CTTTACATA 2220

WO 02/086443

PCT/US02/12476

AAATCAAGGA GACTGCTGAA AATCTCTAAG GGACAGGATT TTCCAGATCC TAATTGGAAA 2280
TTTAGCAATA AGGAGAGGAG TCCAAGGGGA CAAATAAAGG CAGAGAGAGA GAGAGAGAGA 2340
GGGAGAGGAA GAAAGAGAG AGAGAAAAGA GCCTCGTGCC

Seq ID NO: 429 Protein sequence
Protein Accession #: NP_003705

1 11 21 31 41 51
MCAERLGQFM TLALVLATFD PARGTDATNP PEGPQDRSSQ QKGRSLQNT AEIQHCLVNA 60
GDVGGGVFEC FENNCEIRG LHGICMTFLH NAGKFDAGGK SFIKDALKCK AHALRHRFGC 120
ISRKCPAIRE MVSQIQRECY LKHDLCAAAQ ENTRVIVEMI HFKDLLLHEP YVDLVNLLLT 180
QGEVKEAIT HSVQVQCEQN WSLCSILSF CTSALQKPPT APPERQPPVD RTKLSRAHHG 240
EAGHHLPEPS SRBTGRGAKG ERGSKSHPNA HARGRVGGLG AQGPSGSSEW EDEQSEYSDI 300
RR

Seq ID NO: 430 DNA sequence
Nucleic Acid Accession #: NM_005940
Coding sequence: 23..1489

1 11 21 31 41 51
AAGCCAGCA GCCCGGGGGG GGATGGCTCC GGCGGCTGG CTCCGAGCG CGGCCGCGCG 60
CGCCCTCTCG CCCCAGATGC TGCTGCTGCT GCTCCAGCG CGCCGCTGCG TGGCCCGGGC 120
TCTGCGCGCG GACGTCCACC ACCTCCATGC CGAGAGGAG GGGCCACAGC CTGGCATGC 180
AGCCCTGCCC AGTAGCCCGG CACCTGCCCC TGCCACGCG GAAGCCCCC GCCTGCGCAG 240
CAGCCTCAGG CCTCCCGCT GTGGCGTGCC CGACCATCT GATGGGCTGA GTGCCCGCA 300
CCGACAGAA AGGTTCGTGC TTTCTGGCGG GCGCTGGGAG AAGACGGACC TCACCTACAG 360
GATCCTTCGG TCCCATGCG AGTTGGTGCA GGAGCAGGT CGGCAGACGA TGGCAGAGGC 420
CCTAAAGTGA TGGAGCGATG TGACGCCACT CACCTTTACT GAGGTGCAAG AGGGCCGTGC 480
TGACATCATG ATCGACTTCG CCAGGTACTG GCATGGGGAC GACCTGCCGT TTGATGGGCC 540
TGGGGGCATC CTGGCCCATG CCTTCTTCCC CAAGACTCAC CGAGAAGGGG ATGTCCACTT 600
CGACTATGAT GAGACCTGGA CTATCGGGGA TGACACGGGC ACAGACCTGC TGCAGGTGGC 660
AGCCCATGAA TTTGGCCAGG TGCTGGGGCT GCAGCACACA ACAGCAGCCA AGGCCCTGAT 720
GTCCGCTTTC TACACTTTC GCTACCCACT GAGTCTCAGC CCAGATGACT GCAGGGCGGT 780
TCAACACCTA TATGGCCAGC CCTGGCCAC TGTCACCTCC AGGACCCAG CCCTGGGCC 840
CCAGGCTGGG ATGACACCA ATGAGATTGC ACCGCTGGAG CCAGACGCC CGCCAGATGC 900
CTGTGAGGCC TCCTTTGAGC CGGTCTCCAC CATCGAGGC GAGCTCTTTT TCTTCAAGC 960
GGGCTTGTG TGGCGCCTCC GTGGGGGCCA GCTGCAGCCC GGCTACCCAG CATTTGGCTC 1020
TCGACCTGAG CAGGACTGCG CCAGCCCTGT GAGCGCTGCC TTCAGGATG CCCAGGGCCA 1080
CATTTGGTTC TTCCAAAGGT CTCAGTACTG GGTGTACGAC GGTGAAAAGC CAGTCTGGG 1140
CCCCGACCCC CTCACCGAGC TGGGCTGGT GAGGTTCCTG GTCCATGCTG CTTTGGCTG 1200
GGGTCCCGAG AAGAACAGA TCTACTTCTT CCGAGGCAGG GACTACTGGC GTTTCACCC 1260
CAGCACCCGG CGGTAGACA GTCCCGTGCC CCGCAGGGCC ACTGACTGGA GAGGGGTGCC 1320
CTCTGAGATC TACACTTTC TCCAGGATGC TGATGGCTAT GCCTACTTCC TCGCGGGCG 1380
CCTCTACTGG AAGTTTGACC CTGTGAAGGT GAAGGCTCTG GAAGGCTTCC CCGCTCTCGT 1440
GGGTCTGAC TTTCTTGGCT GTGCCGAGCC TGCCAACTAT TTCTCTGAC CATGCTTGG 1500
ATGCCCTCAG GGGTGTGAC CCCTGCCAGG CCACGAATAT CAGGTAGAG ACCCATGGCC 1560
ATCTTGTGCG CTGTGGGAC CAGGCATGGG ACTGAGCCCA TGTCTCTGCG AGGGGATGG 1620
GGTGGGATC AACACCATG ACAACTGCCG GGAGGGCCAC GCAGTCTGCG GTCACCTGCC 1680
AGCGACTGTC TCAGACTGGG CAGGAGGCT TTGGCATGAC TTAAGAGGAA GGGCAGTCTT 1740
GGGACCGCT ATGCAGGTCC TGGCAAACCT GGCTGCCCTG TCTCATCCCT GTCCCTCAG 1800
GTAGCACCAT GGCAGGACT GGGGAACCTG AGTGTCTTGG CTGTATCCCT GTTGTGAGST 1860
TCCTTCCAGG GGCTGGCAGT GAAGCAAGGG TGCTGGGCCC CCATGGCCTT CAGCCCTGCC 1920
TGAGCAACTG GAGCTAGGG CAGGGCCACT TCCTGAGGTC AGGTCTTGGT AGGTGCTGCT 1980
ATCTGTCTGC CTTCTGGCTG ACAATCTCTG AAATCTGTTT TCCAGAAATC AGGCCAAAAA 2040
GTTACAGTC AAATGGGGAG GGGTATTCCT CATGCAGGAG ACCCCAGGCC CTGGAGGCTG 2100
CAACATACCT CAATCCTGTC CCAGGCCGGA TCCTCTGAA GCCCTTTTCG CAGCACTGCT 2160
ATCCTCCAAA GCCATTGTA ATGTGTGTAC AGTGTGTATA AACCTTCTTC TTTCTTTTTT 2220
TTTTAACT GAGGATTGTC ATTAAACACA GTTGTTTTCT

Seq ID NO: 431 Protein sequence
Protein Accession #: NP_005931

1 11 21 31 41 51
MAPAAWLRS AARALLPPLL LLLIQPPPLL ARALPFDVHH LHAERRGPQP WHAALPSSPA 60
PAPATQEAPR PASSLRPPRC GVPDPDGLS ARNRQKRFVL SGRWEKTDL TYRILRFPWQ 120
LVQEQVRQTM AEALKVWSDV TPLTFTEVHE GRADIMIDFA RYWHGDDLPF DGPGGILAH 180
FFPKTHREGD VHFYDETWI IGDDQGTDL LQVAHEFGHV LGLQHTTAAK ALMSAFYTF 240
YPLSLSPDDC RGVQHLVQGP WPTVTSRTPA LGPQAGIDTN EIAPLEPDAP PDACEASFDA 300
VSTIRGELFF FKAGFVWRLR GQQLQPGYPA LASRHQQLP SPVDAAPEDA QGHINFFQGA 360
QYVVDGEKP VLGPAPLTEL GLVRFPVHAA LVNGPEKNKI YFFRGRDYWR FHPSTRVDS 420
FVPRRATDWR GVPSEIDAAF QDADGYAYFL RGRLYWKFDP VRVKALEGFP RLVGPDFFGC 480
AEPANTFL

Seq ID NO: 432 DNA sequence
Nucleic Acid Accession #: NM_024022
Coding sequence: 202..1563

1 11 21 31 41 51
ACCGGGCACC GGAAGGCTCG GGTACTTTTC TTCTTAATTA GTTCATGCCC GTGTGAGCCA 60
GGAAAGGGCT GTGTTTATGG GAAGCCAGTA ACATGTGGCC TACTATCTCT TTCCTGGTGT 120
CCATCTACAT TTTGGGACT CGGAATTAT GAGGTAGAGG TGGAGGCGGA GCCGATGTC 180
AGAGTCTCTG AAATAGTCAC CATGGGGGAA AATGATCCGC CTGCTGTGTA AGCCCCCTTC 240
TCATTCCGAT CGCTTTTTCG CCTGTATGAT TTGAAAATAA GTCTGTGTGC ACCAGATGCA 300

WO 02/086443

PCT/US02/12476

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

GATGCTGTTG CTGCACAGAT CTTGTCACTG CTGCCATTGA AGTTTTTTTC AATCATCGTC 360
 ATTGGGATCA TTGCATTGAT ATTAGCACTG GCCATTGGTC TGGGCATCCA CTTGACTGTC 420
 TCACGGGAAGT ACAGATGTGCT CTATCTCTTT AAGTGATATG AGCTGATAGC TCGATGTGAC 480
 GGAGTCTCGG ATTGCAAAAG CCGGGAGGAC GAGTACCGCT GTGTCCGGGT GGGTGGTCAG 540
 AATGCCGTGC TCCAGGTGTT CACAGCTGCT TCGTGAAGA CCATGTGCTC CAGTACTGTC 600
 AAGGGTCACT ACGCAAAATG TGCTGTGCC CAATCGGGTT TCCCAAGCTA TGTGAGTTCA 660
 GATAACCTCA GAGTGAAGCT GCTGGAGGGG CAGTTCGGGG AGGAGTTTGT GTCCATCGAT 720
 CACCTCTTGC CAGATGACAA GGTGACTGCA TTACCACTC CAGTATATGT GAGGGAGGGA 780
 TGTGCTCTTG GCCACGTGGT TACCTTGCAG TGCACAGCCT GTGGTCATAG AAGGGGTAC 840
 AGCTCAAGCA TCGTGGGTGG AAACATGTCC TTGCTCTGCG AGTGGCCCTG SCAGGCCAGC 900
 CTTGAGTTCC AGGGCTACCA CTTGTGCGGG GGCTCTGTCA TCACGCCCTT GTGGATCATC 960
 ACTGTGTCAC ACTGTGTTTA TGACTGTGAC CTCCCAAGT CATGGACCAT CCAGGTGGGT 1020
 CTAGTTTCCC TGTGGGACAA TCCAGCCCCA TCCCACTTGG TGGAGAAGAT TGTCTACCAC 1080
 AGCAAGTACA AGCCAAAGAG GCTGGGCAAT GACATCGCCC TTATGAAGCT GGCGGGGCA 1140
 CTCACGTTCA ATGAAATGAT CCAGCCTGTG TGCTGCCCCA AGCTGGAAGA GAACTTCCCC 1200
 GATGGAAGAG TGTGCTGGAC GTGAGGATGG GGGGCCACAG AGGATGGAGG TGACGCCCTCC 1260
 CCTGTCTGTA ACCACGGGGT CGTCCCTTTG ATTTCCAAAC AGATCTGCAA CCACAGGGAC 1320
 GTGTACGGTG GCATCATCTC CCCTCCATG CTCTGCGGGG GCTACCTGAC GGGTGGCGTG 1380
 GACAGCTGCC AGGGGGACAG CCGGGGGGCC CTGCTGTGTC AAGAGAGGAG GCTGTGGAAG 1440
 TTAGTGGGAG CGACCAAGTT TGGCATCGGC TGGCAGAGG TGAACAAGCC TGGGGTGTAC 1500
 ACCCGTGTCA CCTCTTCTCT GGACTGGATC CACGAGCAGA TGGAGAGAGA CCTAAAAACC 1560
 TGAAGAGGAA GGGGACAGCT AGCCACCTGA GTTCTGAGG TGATGAAGAC AGCCCGATCC 1620
 TCCCTGGAC TCCGTTAGT GAACCTGCAC ACAGAGCAGC ACCCTGGAG CTCTGAGTTC 1680
 CGGCACCACT AGCAGGCCCC AAAGAGGCAC CCTTCCATCT GATTCCAGCA CAACCTTCAA 1740
 GCTGCTTTT GTTTTTTGT TTTTGTAGGT GGAGTCTGCG TCTGTGCCCC AGGCTGGAGT 1800
 GCAGTGGGCA AATCCCTGCT CACTGCAAGC TCCGCTTCCC TGGTTCAGC GATTCTCTTG 1860
 CCTCAGCTTC CCGATAGCT GGGACACAG GTGCCCCCA CCACACCCAA CTAATTTTGT 1920
 TATTTTGTAGT AGAGACAGGG TTTCACCATG TTGGCCAGGC TGCTCTAAA CCCCTGACCT 1980
 CAAATGATGT GCCTGCTTCA GCCTCCACCA GTGCTGGGAT TACAGGCATG GGCCACCAAG 2040
 CCTAGCCTCA CGCTCCTTTC TGATCTTAC TAAGAACAAA AGAAGCAGCA ACTTGCAAGG 2100
 GCGGCTTTC CCACTGGTCC ATCTGGTTT CTCTCCAGGG GTCTTGCAA ATTCTGACG 2160
 AGATAAGCAV REGCAGSHV TLQCTACGHR RGYSSRIVGG NMSLLSQWPW QASLQFQGYH 2220
 CCAGCCAGCA AGTCAGAAC TGCACTCACT GCACGTTTC ATCTTAGGG ACCAGAACCA 2280
 AACCCACCTT TTCTACTTCC AAGACTTATT TTCACATGTG GGGAGGTTAA TCTAGGAATG 2340
 ACTCGTTTAA GGCCTATTTC CATGATTTC TTGTAGCAT TGGTGCTTGA CSTATTATTG 2400
 TCCTTTGATT CCAATAATA TGTTCCTTC CCTCAAAAAA AAAAAAAA 2460
 AAAAA

Seq ID NO: 433 Protein sequence
 Protein Accession #: NP_076927

1 11 21 31 41 51
 MGENDPPAVE AFPSFRSLFG LDDLKISPIVA PDADAVAAQI LSLPLKFFP IIVIGITALI 60
 LALAIGLGIH FDCSGKYRCR SSFKCIELIA RCDGVSDCKD GEDEYRCVRV GQNAVLQVP 120
 TAASWKTMCS DDWKGHYANV ACAQLGFPSY VSSDNLRVSS LEGQFREEFV SIDHLLPDDK 180
 VTALHSHVYV REGCAGSHV TLQCTACGHR RGYSSRIVGG NMSLLSQWPW QASLQFQGYH 240
 LCGGSVITPL WIITAAHCYV DLYLPKSWTI QVGLVSLLDN PPSHLVEKI VYHSKYKPKR 300
 LGNDIALMKL AGPLTFNEMI QVCLPNSEB NFPDGVKWCN SGMGATEDGG DASPVLNHAA 360
 VPLISNKN CNHRDVGIIIS PSMLCAGYLT GGVDSCQGDS GGPLVCQERR LMKLVGATSP 420
 GIGCAEVNKP GVTYRVTSLF DWIHEQMERD LKT

Seq ID NO: 434 DNA sequence
 Nucleic Acid Accession #: NM_000493.2
 Coding sequence: 97..2139

1 11 21 31 41 51
 CACCTTCTGC ACTGCTCATC TGGGCAGAGG AAGCTTCAGA AAGCTGCCAA GGCAACATCT 60
 CCAGGAACCTC CCAGCAGGCA GAATCCATCT GAGAATATGC TGCCACAAT ACCCTTTTGT 120
 CTGCTAGTAT CCTTGAACCT GGTTCATGGA GTGTTTTACG CTGAACGATA CCAATAGCCC 180
 ACAGGCATAA AAGGCCCACT ACCCAACACC AAGACACAGT TCTTCATTCC CTACACCATA 240
 AAGAGTAAAG GTATAGCAGT AAGAGGAGAG CAAGGTACTC CTGGTCCACC AGGCCCTGCT 300
 GGACCTCGAG GGCACCCAGG TCCTTCTGGA CCACCAGGAA AACCAGGCTA CGGAAGTCTC 360
 GGACTCCAAG GAGAGCCAGG GTTGCCAGGA CCACCGGAC CATCAGCTGT AGGGAACCA 420
 GGTGTGCCAG GACTCCAGG AAAACAGGA GAGAGAGGAC CATATGGACC AAAAGGAGAT 480
 GTTGGACCACT CTGGCCTACC AGGACCCCGG GGCCACACAG GACCACTGG AATCCCTGGA 540
 CCGGCTGGAA TTTCTGTGCC AGGAAACCTC GGACAACAGG GACCCACAGG AGCCCCAGGA 600
 CCCAGGGGCT TTCTCGGAGA AAAGGGTGCA CCAGGAGTCC CTGGTATGAA TGGACAGAAA 660
 GGGGAAATGG GATATGTTGC TCTGTTGCTG CCAGGTGAGA GGGGTCTTCC AGGCCCTCAG 720
 GGTCCACAG GACCATCTGG CCTCCTGGA GTGGGAAAAA GAGGTGAAAA TGGGGTTCCA 780
 GGACAGCCAG GCATCAAAGG TGATAGAGGT TTTCCGGGAG AAATGGGACC AATTGGCCCA 840
 CCAGGTCCCC AAGGCCCTCC TGGGGAACGA GGGCCAGAA GCAATTGAAA GCCAGGAGCT 900
 GCTGGAGCCC CAGGCCAGCC AGGGATTCCA GGAACAAAG GCTTCCCTGG GGCTCCAGGA 960
 ATAGCTGGGC CCCAGGGGCC TCTGGCTTT GGGAAACAG GCTTGCCAGG CCTGAAGGGA 1020
 GAAGAGGAC CTGCTGGCCT TCTGGGGGT CCAGGTGCCA AAGGGGAACA AGGGCCAGCA 1080
 GGTCTTCTG GGAAGCCAGG TCTGACTGGA CCCCCTGGGA ATATGGGACC CCAAGGACCA 1140
 AAAGGCATCC CGGTAGCCA TGGTCTCCA GGCCCTAAAG GTGAGACAGG GCCAGCTGGG 1200
 CCTGCAGGAT ACCCTGGGGC TAAGGGTGAA AGGGGTTCCC CTGGGTGAGA TGGAAAAACA 1260
 GGGTACCCAG GAAACACAGG TCTCGATGTT CTAAGGGTA ACCCAGGGTT ACCAGGTCCA 1320
 AAAGGTGATC CTGAGTTGG AGGACCTCCT GGTCTCCAG GCCCTGTGGG CCCAGCAGGA 1380
 GCAAAGGAA TGCCCGACA CAATGAGAG GCTGGCCCA GAGGTGCCCC TGGAAATACA 1440
 GGTACTAGAG GGCCTATTGG GGCACAGGC ATTCCAGGAT TCCCTGGGTC TAAAGGGGAT 1500
 CCAGGAAGTC CCGTCTCTCC TGGCCAGCT GGCATAGCAA CTAAGGGCCT CAATGGACCC 1560
 ACCGGGCCAC CAGGGCTTCC AGGTCCAAGA GGCCACTCTG GAGAGCCTGG TCTTCCAGGG 1620
 CCCCCTGGGC CTCCAGGCC ACCAGGTCAA GCAGTATGCG CTGAGGGTTT TATAAGGCA 1680
 GGCCAAAGGC CCACTCTTTC TGGGACCCCT CTGTGATGTC CCAACAGGG GGTAAACAGGA 1740

5 ATGCCCTGTG CTGCTTTTAC TGTATTCTC TCCAAAGCTT ACCCAGCAAT AGGAACTCCC 1800
 ATACCATTG ATAAAAATTT GTATAACAGG CAACAGCATT ATGACCCAAG GACTGGAATC 1860
 TTTACTGTG AGATAACAGG AATATACTAT TTTTCATACC ACGTGCAATG GAAAGGGACT 1920
 CATGTTTGGG TAGGCGCTGT TAAGAATGGC ACCCTGTAA TGTACACCTA TGATGAATAC 1980
 ACCAAAGGCT ACCTGGATCA GCCTTCAGGG AGTGCCATCA TCGATCTCAC AGAAAAATGAC 2040
 CAGGTGTGGC TCCAGCTTCC CAATGCCGAG TCAATGGCC TATACTCTC TGAGTATGTC 2100
 CACTCTCTT TCTCAGGATT CCTAGTGGCT CCAATGTGAG TACACCCAC AGAGCTAATC 2160
 TAAATCTGT GCTGAAAAA GCATTCTCTA ACTCTACCCC ACCCTACAAA ATGCATATGG 2220
 10 AGGTAGGCTG AAAAGAAATG AATTTTTATT TTCTGAAATA CAGATTGAG CTATCAGACC 2280
 AACAAACCTT CCCCTGAAA AGTGAGCAGC AACGTAATA CGTATGTGAA GCCTCTCTTG 2340
 AATTTCTAGT TAGCAATCTT AAGGCTCTTT AAGGTTTTCT CCAATATTAA AAAATATCAC 2400
 CAAAGAACTG CTGCTATGTT AAAAACAAAC AACAAAAAAC AAAGCAACAA AAAAAAAT 2460
 TAAAAAATA AACAGAAATA GAGCTCTAAG TTATGTGAAA TTGATTGTA GAAACTCGGC 2520
 15 ATTTCTTTT TAAAGAGGCC TGTCTTAAC ATGAATATG AGAACTCTA GGAACATCC 2580
 AGGAGGTATC ATATAACTTT GTAGAACTTA AATACTTGA TATTCAAAT TAAAGACAC 2640
 TGTATCCCTT AAAATATTT TGATGGTGA CTACTCTGAG GCCTGTATGG CCCCTTTCAT 2700
 CAATATCTAT TCAAAATATC AGGTGCATAT ATACTTGTTA AAGCTCTTAT ATAAAAAGC 2760
 CCCAAATAT TGAAGTTCAT CTGAAATGCA AGGTGCTTTC ATCAATGAAC CTTTTCAAAA 2820
 20 CTTTCTATG ATTGCAGAGA AGCTTTTTAT ATACCCAGCA TAACTTGGA ACAGGTATCT 2880
 GACCTATTCT TATTAGTGA ACACAAGTGT GATTAAATTT ATTTCTTAA TTCTTATTG 2940
 AATCTTATG GATATGATTT TCTGGATTTA CAGAACATTA GCACATGTAC CTGTGCGCTC 3000
 CCATTCAAGT GAAGTTATAA TTTACACTGA GGGTTTCAA ATTGACTAG AAGTGGAGAT 3060
 ATATTATTA TTTATGCACT GTACTGTATT TTTATATTGC TGTTTAAAA TTTTAAAGCTG 3120
 25 TGCCCTCACT ATTAAGACAC AAAATGTTTT ACCTACTCCT TATTATCGAC ACAATAAAAT 3180
 AACATCAATA GATTTTTAGG CTGAATTAAT TTGAAGCAG CAATTGTCTG TTCTCAACCA 3240
 TTCTTCAAG GCTTTTCATT CGACACAATA AATAACATC AATAG

Seq ID NO: 435 Protein sequence

Protein Accession #: NP_000484.2

30 1 11 21 31 41 51
 MLQPFPFLLL VSLNLVHGVF YAERYQMPTG IKGPLPNTKT QFFIPYTIKS KGIIVRGEQG 60
 TPGPPGAPG RGHPPGSPGP GKPGYGSPLG QGEPGLPGPP GPSAVGKPGV PGLPGKPGER 120
 35 GPYGPBKDVG PAGLPGPRGP PGPPGIPGPA GISVPGKPGQ QGPTGAPGPR GFPPEKGAPG 180
 VPGMNGQKGE MGYGAPGRGP ERGLPGPQGP TGPSGPPGVG KRGENCVPQG PGLKGRGFP 240
 GEMGPIGPPG PQGPPGERGP EGIGKPGAAG APGQPIGPT KGLPGAPGIA GPPGPPGFGK 300
 PGLPGLKGER GPAGLPGGPG AKGEQGPAGL PGKPLGTGPP GNMGPQGPKG IPGSHGLFPG 360
 40 KGETGPAGPA GYPGAKGERG SPGSDGKPGY PGKPLDGPK GNFGLPGPKG DPGVGGPPLG 420
 PGFVGPAGAK GMPGNGEAG PRGAPGIPGT RGPFGPGIP GFPGSKGDPG SPGPPGPAGI 480
 ATKGLNGPTG PPGPPGPRGH SGEPGLPGPP GPPGPPGQAV MPEGFIKAG RPSSLGTPLV 540
 SANQGVITGM VSAFTVILSK AYPAGITPIP FDKILYNRQ HYDPRTGIFT CQIPGIYFYS 600
 YHVHVKGTHV WVGLYKNGTP VMYTYDEYTK GYLDQASGSA IIDLTENDQV WLQLPNAESN 660
 GLYSSEYVHS SFSGLVAPM

Seq ID NO: 436 DNA sequence

Nucleic Acid Accession #: XM_062811

Coding sequence: 1..888

50 1 11 21 31 41 51
 ATGTGGGGCG CTCGCCGCTC GTCCGTCTCC TCATCTGGA ACCGCGCTTC GCTCTGCGAG 60
 CTGCTGTGCG CTGCGCTGCT GGCGGCGGGG GCGAGGGCCA GCGGCGAGTA CTGCCACGGC 120
 TGGCTGGACG CGCAGGGCGT CTGGCGCATC GGCTTCCAGT GTCCGAGCG CTTCGACGGC 180
 55 GCGACGCCA CCATCTGCTG CGGCAGCTGC GCGTTGCGCT ACTGCTGCTC CAGCGCCGAG 240
 GCGCGCTCG ACCAGGCGCG CTGCGACAAT GACCGCCAGC AGGGGCTGG CGAGCCTGGC 300
 CGGGCGGACA AAGACGGCCC CGACGGCTCG GCAGTGCCCA TCTAGTGGC GTTCTCTATT 360
 GTTGGCTCCG TGTTTGTGCG CTTTATCATC TTGGGGTCCC TGGTGGCAGC CTGTGTGCTC 420
 60 AGATGTCTCC GGTCTAAGCA GGATCCCCAG CAGAGCCGAG CCCCAGGGGG TAACCGCTTG 480
 ATGAGACCA TCCCATGAT CCCAGTGCC AGCACTCCC GGGGGTGGT CTCACGCCAG 540
 TCCAGCACAG CTGCCAGTTC CAGCTCCAGC GCCAATCAG GGGCGCGGC GCCCCACA 600
 AGGTCACAGA CCAACTGTTG CTTGCCGGA GGGACCATGA ACAAGTGTA TGTCAACATG 660
 CCCACGAATT TCTCTGTGCT GAAGTGTGAG CAGGCCACCC AGATTGTGCC ACATCAAGGG 720
 65 CAGTATCTGC ATCCCCATA CGTGGGTGAC ACGGTGCAGC ACGACTCTGT GCCCATGACA 780
 GCTGTGCCAC CTTTCATGGA CGGCTGCAG CCTGGTACA GGCAGATTCA GTCCCTCTC 840
 CCTCACACCA ACAGTGAACA GAAGATGTAC CCAGCGGTGA CTGTATAA

Seq ID NO: 437 Protein sequence

Protein Accession #: XP_062811

70 1 11 21 31 41 51
 MWGARRSSVS SSWNAASLLQ LLLAALLAAG ARASGEYCHG WLDAQGVWRI GFQCFERFDG 60
 GDATICGSSC ALRYCCSSAE ARLDQGGCDN DRQQGAGEPG RADKDPDGS AVPIYVPFLI 120
 75 VGSVFVAFII LGLSVAACCC RCLRPKQDPQ QSRAPGNNRL METIFMIPSA STSRGSSSRQ 180
 SSTAASSSSS ANSGARAPPT RSQTNCLLPE GTMNNVYVNM PTNFSVLNQM QATQIVPHQG 240
 QYLHPPFVGY TVQHDSVPMT AVPPFMDGLQ PGYRQIQSPF PHTNSEQKMY PAVTV

Seq ID NO: 438 DNA sequence

Nucleic Acid Accession #: NM_004004.1

Coding sequence: 1..681

80 1 11 21 31 41 51
 85 ATGGATTGGG GCACGCTGCA GACGATCTG GGGGGTGTGA ACAACACTC CACCAGCATT 60
 GGAAAGATCT GGCTCACCCT CCTCTTCAAT TTTGCAATTA TGATCTCTGT TGTGCTGCA 120
 AAGGAGGTGT GGGGAGATGA GCAGGCGGAC TTTGTCTGCA ACACCTGCA GCCAGGCTGC 180

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

```
AAGAACGTGT GCTACGATCA CTACTTCCCC ATCTCCACA TCGGCTATG GGCCTGCG 240
CTGATCTTGG TGTCAGGCC AGCGCTCCTA GTGGCCATGC ACGTGGCCTA CCGGAGACAT 300
GAGAAGAAGA GGAAGTTTCAT CAAGGGGGAG ATAAAGAGTG AATTTAAGGA CATCGAGGAG 360
ATCAAAACCC AGAAGGTCCG CATCGAAGGC TCCCTGTGGT GGACCTACAC AAGCAGCATC 420
TTCTTCCGGG TCATCTTCGA AGCGGCTTC ATGTACGTCT TCTATGTCAT GTACGACGGC 480
TTCTCCATCG AGCGGCTGGT GAAGTGCAAC GCGTGGCCTT GTCCCAACAC TGTGGAGTCG 540
TTTGTGTCCC GGCCACGGGA GAAGACTGTC TTCACAGTGT TCATGATTGC AGTGTCTGGA 600
ATTGTGATCC TGCTGAATGT CACTGAATTG TGTATTTCG TAATTAGATA TTGTTCTGGG 660
AAGTCAAAA AGCCAGTTTA A
```

Seq ID NO: 439 Protein sequence
Protein Accession #: NP_003995.1

1 11 21 31 41 51
MDWGTQLTIL GGVNKHSTSI GKILWLVLF I FRIMILVVA KEVWDEQAD FVCNTLQPGC 60
KNVCYDHYFP ISHRLWALQ LIFVSSPALL VAMHVAYRRH EKRRKFKGE I KSEFKDIEE 120
IKTKQVRIEG SLWWTYTSSI FFRVIFPAAP MYFVVMYDG FSMQRLVKCN ANPCPNTVDC 180
FVSRPTEKTV FTVFMIAVSG ICILNLVTEL CYLLIRYCSG KSKKPV

Seq ID NO: 440 DNA sequence
Nucleic Acid Accession #: XM_061091.1
Coding sequence: 1..2481

1 11 21 31 41 51
ATGCCAAATA CTTCAGGAAC AACAGGATT GAAATTTGGC TTCTCCAAGA GCCGCCGGG 60
CACCGAGCGG TGGTCGCGCG TCTCCTCCG GTGAGTCCCA GCCCGAGTTC GGCTCTGGCG 120
CCCGGGTACC GCGCAGTGCC GGCTGCCGAT GACCGATTCA CGCTCCCGAT GATTGGAGGT 180
CAGATGKATG GTGAGAAAGT AGATCTCTGG AGCCTTGGTG TTCTTTGCTA TGAATTTTAA 240
GTTGGGAAGC TCCTTTTGA GGCAACGAA GTCCATGTAA GCAAGAAAC CATCGGGAAG 300
ATTTCACTG CCAGCAAAAT GATGTGGTGC TCGGCTGCAG TGGACATCAT GTTCTCTGTA 360
GATGGGTCTA ACAGCGTCGG GAAAGGGAGC TTTGAAAGGT CCAAGCACIT TCCATCACA 420
GTCTGTGACG GTCTGGACAT CAGCCCCGAG AGGGTCAGAG TGGGAGCAIT CCAGTTCAGT 480
TCCACTCCTC ATCTGGAATT CCCCTTGGAT TCATTTTCAA CCAACAGGA AGTGAAGGCA 540
AGAATCAAGA GATGTGTTTT CAAAGGAGGG GCGACGGAGA CGGAAGTTC TCTGAAATAC 600
CTTCTGCACA GAGGGTGTCC TGGAGGCAGA AATGCTTCTG TGCCCCAGAT CCTCATCATC 660
GTCAGTGATG GGAAGTCCCA GGGGATGTG GCACTGCCAT CCAAGCAGCT GAAGAAAGG 720
GGTGTCACTG TGTGTGCTGT GGGGTCCAGG TTTCCAGGT GGGAGGAGCT GCATGCAGTG 780
GCCAGCGAGC CTAGAGGGCA GCAGTGTCTG TTGGCTGAGC AGGTGGAGGA TGCCACCAAC 840
GGCCTCTTCA GCACCCCTCAG CAGCTCGGCC ATCTGTCCA GCGCCACGCG AGCTGGGAGC 900
CCCGAGCTTG TCTTCACTGA GCGTTAATG GGCATCTCTC TGATAGGCCC CTGTGACTCG 960
CAGCCCTGCC AGAATGGAGG CACATGTGTT CCAGAAGGAC TGGACGGCTA CCAGTGCCTC 1020
TGCCCGCTGG CCTTTGGAGG GGAGGCTAAC TGTGCCCTGA AGCTGAGCCT GGAATGCAGG 1080
GTGCACTCTC TCTTCTGTCT GGACAGCTCT GCGGGCACCA CTCTGGACGG CTTCCTCGGG 1140
GCCAAAGTCT TCGTGAAGCG GTTGTGCGG GCGGTGCTGA GCGAGGACTC TCGGGGCCGA 1200
GTGGGTGTGG CCACATACAG CAGGGAGCTG CTGGTGGCGG TGCCGTGGGG GGAGTACCA 1260
GATGTGCTTG ACCTGGTCTG GAGCCTCGAT GGCATTCCCT TCGGTGGTGG CCCCACCTG 1320
ACGGGCAAGT CCTTGGCGGA GCGGCGAGAG CGTGCTTTCG GGAGCGCCAC CAGGACAGGC 1380
CAGGACCGGC CACGTAGAGT GGTGGTTTTC CTCACTGAGT CACACTCCGA GGATGAGGTT 1440
GCGGGCCCGC GCGGTCAAGC AAGGGCGCGA GAGCTGCTCC TGCTGGGTGT AGGCAGTG 1500
GCCGTGCGGG CAGAGCTGGA GGAGATCACA GGCAGCCCAA AGCATGTGAT GGTCTACTCG 1560
GATCCTCAGG ATCTGTTCAA CCAATCCCT GAGCTGCAGG GGAAGCTGTG CAGCCGCGAG 1620
CGGCCAGGGT GCGGACACA AGCCCTGGAC TCGCTCTTCA TGTGGACAC CTCTGCCTCA 1680
GTAGGGCCCG AGAATTTTGC TCAGATGCAG AGCTTTGTGA GAAGCTGTGC CCTCCAGTT 1740
GAGGTGAAC CTGACGTGAC ACAGGTCCGC CTGGTGGTGT ATGGCAGCCA GGTGCAGACT 1800
GCCTTCGGGC TGGACACCAA ACCCACCGG GCTGCGATGC TGCGGGCCAT TAGCCAGGCC 1860
CCCTACCTAG GTGGGGTGGG CTCAGCCGGC ACCGCCCTGC TGCACATCTA TGACAAAGT 1920
ATGACCGTCC AGAGGGGTGC CCGGCTGGT GTCCCCAAG CTGTGGTGGT GCTCACAGGC 1980
GGGAGAGCGG CAGAGGATGC AGCCGTTCCT GCCCAGAAGC TGAGGAACAA TGGCATCTCT 2040
GTCTTGTGTC TGGCGTGGG GCCTGTCTTA AGTGAGGGTC TCCGAGGCT TGCAGGTCCC 2100
CGGATTTCCC TGATCCACGT GCAGCTTAC GCGACCTGC GGTACCACA GGACGTGCTC 2160
ATTGAGTGGC TGTGTGGAGA AGCCAAGCAG CCAGTCAACC TCTGCAACC CAGCCCGTGC 2220
ATGAATGAGG GCAGCTGCGT CCTGCAGAA GGGAGCTACC GCTGCAAGT TCGGGATGGC 2280
TGGGAGGGCC CCCACTGCGA GAACCGTGAG TGGAGCTCTT GCTCTGTATG TGTGAGCCAG 2340
GGATGTGATT TTGAGACGCC CCGTGGGCAC ATGGCTCCCG TGCAGGAGG CAGCAGCGT 2400
ACCCCTCCCA GCAACTACAG AGAAGGCGTG GGCAGTAAA TGGTGCTTAC CTTCTGGAAT 2460
GTCTGTGCCC CAGGTCTCTA G

Seq ID NO: 441 Protein sequence
Protein Accession #: XP_061091.1

1 11 21 31 41 51
MPNTSGTTRI EIWLQEPFG HRLVAALLP VSPSPALALA PGYPPVPAAD DRPTLFMIGG 60
QMHEKVDLM SLGLVCEYFL VGKPPFEANE VHSVKEITGK ISAASKMMWC SAAVDIMPLL 120
DGSNSVKGKS FERSKHFAIT VCDGLDISPE RVRVGAFQFS STPHLEPLD SFSTQOEKKA 180
RIKRMVFKGG RTETELALKY LLHRLPGGR NASVPQILII VTDGKSQGDV ALPSKQLKER 240
GVTVFVGVVR FPRWEELHAL ASEPRGQHLV LAEQVEDATN GLFSTLSSSA ICSSATPAGS 300
PELVFMERLM GISLIGPCDS QPCQNGGTCV PEGLDGYQCL CPLAFGEAN CALKLSLECR 360
VDLLFLDSS AGTTLDFGLR AKVFKRFVR AVLSEDSRAR VGVATYSREL LVAVPVGEYQ 420
DVPDLVWSLD GIPFRGGPTL TGSALRQAAE RGFGSATRTG QDRPRRVVVL LTESHSEDEV 480
AGPARHARAR ELLLLGVGSE AVRAELEIIT GSPKHVMVYS DPQDLFNQIP ELQKLCRSQ 540
RPGCRTQALD LVFPLDTSAS VGPFNAQMQ SFVRSALQF EVNPDVTQVG LVVYGSQVQT 600
AFGLDTKPTR AMLRAISQA PYLGGVGSAG TALLHIYDKV MTVQRGARPG VPKAVVVLGT 660
GRGAEDAAVP AQKLNRNNG VLVVGVGPVL SEGLRRLAGP RDSLIHVAAY ADLRYHQDVL 720

IEWLCEAKQ PVNLCKPSPC MNEGSCVLQN GSYRCKCRDG WEGPHCENRE WSSCSVCVSQ 780
GWILETFLRH MAPVQEGSSR TPPSNYREGL GTEMVPTFWN VCAPGP

Seq ID NO: 442 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..2424

```

1      11      21      31      41      51
10  |-----|-----|-----|-----|-----|
    |ATGCCCCCTT|TCCTGTTGCT|GGAGGCCGTC|TGTGTTTCC|TGTTTTCCAG|AGTGCCCCCA| 60
    |TCTCTCCCTC|TCCAGGAAGT|CCATGTAAGC|AAAGAAACCA|TCGGGAAGAT|TTCAGCTGCC| 120
    |AGCAAAATGA|TGTGGTGCTC|GGCTGCAGTG|GACATCATGT|TTCTGTTAGA|TGGGTCTAAC| 180
    |AGCGTCGGGA|AAGGGAGCTT|TGAAAGGTCC|AAGCACTTTG|CCATCAGAGT|CTGTGACGGT| 240
    |CTGGACATCA|GCCCGAGAGG|GGTCAGAGTG|GGAGCATTTCC|AGTTCAGTTC|CACTCCTCAT| 300
    15 |CTGGAATTCC|CCTTGGATTC|ATTTTCAACC|CAACAGGAAG|TGAAGGCAAG|AATCAAGAGG| 360
    |ATGTTTTCCT|AAGAGAGGGC|CACGGAGACG|GAACTTGCTC|TGAATACCT|TCTGCACAGA| 420
    |GGGTTCGCTG|GAGGCAGAAA|TGCTTCTGTG|CCCCAGATCC|TCATCATCGT|CACTGATGSG| 480
    |AAGTCCACAG|GGGATGTGGC|ACTGCCATCC|AAGCAGCTGA|AGGAAAGGGG|TGTCACTGTG| 540
    |TTTGCTGTGG|GGGTCAAGTT|TCCCAGGTGG|GAGGAGCTGC|ATGCACTGGC|CAGCGAGCCT| 600
    20 |AGAGGGCAGC|ACGTGCTGTT|GGCTGAGCAG|GTGGAGGATG|CCAACAACGG|CCTCTTCAGC| 660
    |ACCCTCAGCA|GCTCGGCCAT|CTGCTCCAGC|GCCACGCCAG|ACTGCAGGGT|CGAGGCTCAC| 720
    |CCCTGTGAGC|ACAGGACGCT|GGAGATGGTC|CGGAGGTTGG|CTGGCAATGC|CCCATGCTGG| 780
    |AGAGGATGCG|GGCGGACCTC|TGCGGTGCTG|GCTGCACACT|GTCCCTTCTA|CAGCTGGAAG| 840
    |AGAGTGTTC|TAAACCAACC|TGCCACCTGC|TACAGGACCA|CCTGCCCAGG|CCCCGTGTAC| 900
    25 |TCGCAGCCCT|GCCAGAATGG|AGGCACATGT|GTCCAGAAG|GACTGGACGG|CTACCACTGC| 960
    |CTCTGCCCGC|TGGCCTTTGG|AGGGGAGGCT|AACTGTGCCC|TGAAGCTGAG|CCTGGAATGC| 1020
    |AGGGTCGACC|TCCTCTTCCT|GCTGGACAGC|TCTGCGGGCA|CCACTCTGGA|CGGCTTCCTG| 1080
    |AGGCCCAGAG|GCTTCGTGTC|GCGGTTTGTG|CGGGCCGTGC|TGAGCGAGGA|CTCTCGGGCC| 1140
    |CGAGTGGGTG|TGGCCACATA|CAGCAGGGAG|CTGCTGTGGG|CGGTGCCTGT|GGGGGAGTAC| 1200
    30 |CAGGATGTGC|CTGACCTGGT|CTGGAGCCTC|GATGGCATTC|CCTTCCGTGG|TGGCCCCACC| 1260
    |CTGACCGGCA|GTGCCCTTGC|GCAGGCGGCA|GAGCGTGGCT|TCGGGAGCGC|CACCAGGACA| 1320
    |GGCCAGGACC|GGCCACGTAG|AGTGTGTGTT|TTGCTCACTG|AGTCACACTC|CGAGGATGAG| 1380
    |GTTGCGGGCC|CAGCGCTGCA|CGCAAGGGCG|CGAGAGCTGC|TCCTGCTGGT|TGTAGGCAGT| 1440
    |GAGGCCGTGC|GGGCAGAGCT|GGAGGAGATC|ACAGGACGCC|CAAAGCATGT|GATGCTCTAC| 1500
    35 |TCGGATCCTC|AGGATCTGTT|CAACCAATC|CCTGAGCTGC|AGGGGAAGCT|GTGCAGCCGG| 1560
    |CAGCGGCCAG|GGTGCCCGAC|ACAAGCCCTG|GACCTCGTCT|TCATGTTGGA|CACCTCTGCC| 1620
    |TCAGTAGGGC|CCGAGAATTT|TGCTCAGATG|CAGAGCTTTG|TGAGAAGCTG|TGCCCTCCAG| 1680
    |TTTGAGGTGA|ACCTGTAGCT|GACACAGGTC|GGCCTGGTGG|TGTATGGCAG|CCAGGTGCAG| 1740
    |ACTGCCCTTC|GGCTGGACAC|CAAAACCCACC|CGGGCTGCGA|TGCTGCGGGC|CATTAGCCAG| 1800
    40 |GCCCCCTACC|TAGGTGGGGT|GGGCTCAGCC|GGCACCGCCC|TGCTGCACAT|CTATGACAAA| 1860
    |GTGATGACCG|TCCAGAGGGG|TGCCCGGCCT|GGTGTCCCCA|AAGCTGTGGT|GGTGCTCACA| 1920
    |GGCGGGAGAG|GCGCAGAGGA|TGACGCCGTT|CTGCGCCAGA|AGCTGAGGAA|CAATGGCATC| 1980
    |TCTGTCTTGG|TCGTGGGGCT|GGGGCCTGTC|CTAAGTGAGG|GTCTGCGGAG|GCTTGCAAGT| 2040
    |CCCCGGGATT|CCCTGATCCA|CGTGGCAGCT|TACGCCGACC|TGCGGTACCA|CCAGGACGTG| 2100
    45 |CTCATTGAGT|GGCTGTGTGG|AGAAGCCAAG|CAGCCAGTCA|ACCTCTGCAA|ACCCAGCCCG| 2160
    |TGCAATGAAT|AGGGCAGCTG|CGTCTGCAG|AATGGGAGCT|ACCCTGCAA|GTGTCGGGAT| 2220
    |GGCTGGGAGG|GCCCCCACTG|CGAGAACCGT|GAGTGGAGCT|CTTGCTCTGT|ATGTGTGAGC| 2280
    |CAGGGATGGA|TTCTTGAGAG|GCCCTGAGG|CACATGGCTC|CCGTGCAGGA|GGGCAGCAGC| 2340
    50 |CGTACCCCTC|CCAGCAACTA|CAGAGAAGGC|CTGGGCACTG|AAATGGTGCC|TACCTTCTGG| 2400
    |AATGTCGTGG|CCCCAGGTCC|TTAG
  
```

Seq ID NO: 443 Protein sequence
Protein Accession #: Eos sequence

```

1      11      21      31      41      51
55 |-----|-----|-----|-----|-----|
    |MPPFLLEAV|CVFLFSRVPP|SLPLQEVHVS|KETIGKISAA|SKMMWCSAAV|DIMFLLDGSN| 60
    |SVGKGSFERS|KHFAITVCDG|LDISPERVRV|GAFQFSSTPH|LEPPLDSFST|QQEVKARIKR| 120
    |MVFKGGRTET|ELALKYLLHR|GLPGGRNASV|PQILIIIVTD|KSGQDVALPS|KQLKERGVTV| 180
    60 |FAVGVRPPRW|EELHALASEP|RGQHVLLAEQ|VEDATNGLFS|TLSSSAICSS|ATPDCRVEAH| 240
    |PCEHRTLEMY|REFAGNAPCW|RGSRRTLAVL|AAHCPFYSWK|RVFLTHPATC|YRTTCGPGCD| 300
    |SQPCQNGGTC|VPEGLDSYQC|LCPLAFGGEA|NCALKLSLEC|RVDDLFLDLS|SAGTTLDGFL| 360
    |RAKVVFVRFY|RAVLSDSRA|RVGVATYSRE|LLVAVPVGEY|QDVPLVWVSL|DGIPFRGGPT| 420
    |LTGSALRQAA|ERGFSGSATRT|GQDRPRRVVV|LLTESHSEDE|VAGPARHARA|RELLLLGVGS| 480
    65 |EAVRAELEEI|TGSPKHMVMY|SDPQDLFNQI|PELQKLCRSR|QRPGRCTQAL|DLVFMIDTSA| 540
    |SVGPENFAQM|QSFVRSCALQ|FEVNPDTVQV|GLVVYGSQVQ|TAFGLDTKPT|RAAMLRAISQ| 600
    |APYLGGVGSA|GTALLHIYDK|VMTVQRGARP|GVPKAVVVLV|GGRGAEDAAV|PAQLKRNNGI| 660
    |SVLVVGVGPF|LSEGLRLLAG|PRDSLIIHVA|YADLRYHQDV|LIEWLCEAK|QPVNLCKPSP| 720
    |CMNEGSCVLQ|NGSYRCKCRD|WNEGPHCENR|ENSSCSVCVS|QGWILETFLR|HMAPVQEGSS| 780
    70 |RTPPSNYREG|LGTEMVPTFW|NVCAPGP
  
```

Seq ID NO: 444 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 89..2356

```

1      11      21      31      41      51
75 |-----|-----|-----|-----|-----|
    |GCCCCCTGGC|CCGAGCCGCG|CCCCGGTCTG|TGAGTAGAGC|CGCCCGGGCA|CCGAGCGCTG| 60
    |GTCGCGGCTC|TCCTTCCGTT|ATATCAACAT|GCCCCCTTTC|CTGTTGCTGG|AAGCCGCTCT| 120
    80 |TGTTTTCTCG|TTTTCAGAG|TGCCCCCATC|TCTCCCTCTC|CAGGAAGTCC|ATGTAAGCAA| 180
    |AGAAACCATC|GGGAAGATTT|CAGCTGCCAG|CAAAATGATG|TGGTGCTCGG|CTGCAGTGGG| 240
    |CATCATGTTT|CTGTAGATG|GGTCTAACAG|CGTCGGGAAA|GGGAGCTTTG|AAAGTCCCAA| 300
    |GCACCTTGGC|ATCACAGTCT|GTGACGGTCT|GGACATCAGC|CCGAGAGGGG|TCAGAGTGGG| 360
    |AGCATTCCAG|TTCAGTTCCA|CTCCTCATCT|GGAATTCGCC|TTGATTTCAT|TTTCAACCCA| 420
    85 |ACAGGAAGTG|AAGGCAAGAA|TCAAGAGGAT|GGTTTTCAAA|GGAGGGCGCA|CGGAGACGGA| 480
    |ACTTGCTCTG|AAATACCTTC|TGCAAGAGG|GTTGCTCGGA|GGCAGAAATG|CTTCTGTGCC| 540
    |CCAGATCCTC|ATCATCTGCA|CTGATGGGAA|GTCCCAGGGG|GATGTGGCAC|TGCCATCCAA| 600
  
```

5	GCAGCTGAAG	GAAGGGGTG	TCACGTGTT	TGCTGTGGG	GTGAGTTTC	CCAGGTGGGA	660
	GGAGCTGCAT	GCATGCGCA	GCGAGCTAG	AGGSCAGCAC	GTGCTGTGG	CTGAGCAGGT	720
	GGAGATGCC	ACCAACGGCC	TCTTCAGCAC	CCTCAGCAGC	TGCGCATCT	GCTCCAGGCC	780
	CACGCCAGAC	TGCAGGGTCG	AGGCTCACCC	CTGTGAGCAC	AGGAGCTGG	AGATGGTCCG	840
	GGAGTTGCT	GGCAATGCC	CATGCTGGAG	AGGATCGGG	CGGACCCCTG	CGGTGCTGGC	900
	TGCACACTGT	CCCTTCTACA	GCTGGAAGAG	AGTGTCTCA	ACCCACCCTG	CCACCTGTCTA	960
	CAGGACCAAC	TGCCAGGGC	CCTGTGACTC	GCAGCCCTGC	CAGAATGGAG	GCACATGTGT	1020
	TCCAGAAGGA	CTGGACGGCT	ACCAGTGCCT	CTGCCCGCTG	GCCTTTGAG	GGGAGGCTAA	1080
10	CTGTGCCCTG	AAGCTGAGCC	TGGAATGCAG	GGTCGACCTC	CTCTTCCTGC	TGGACAGCTC	1140
	TGCGGGCAAC	ACTCTGGAAG	GCTTCTGCG	GGCCAAAGTC	TTGCTGAAGC	GGTTTGTGGG	1200
	GGCGGTGCTG	AGCGAGGACT	CTCGGGCCCG	AGTGGGTGTG	GCCACATACA	GCAGGGAGCT	1260
	GCTGGTGGCG	GTGCGTGTGG	GGGAGTACCA	GGATGTGCCT	GACCTGGTCT	GGAGCCTCGA	1320
	TGGCATTCCC	TTCCGTGGTG	GCCCCACCTT	GACGGGCAGT	GCCTTGGCGC	AGGCGGCAGA	1380
	GGGTGGCTTC	GGGAGCGCCA	CCAGGACAGG	CCAGGACCGG	CCACGTAGAG	TGGTGGTTTT	1440
15	GCTCATGAG	TCGACATCCG	AGGATGAGGT	TGCGGGCCCA	GCGCGTCACG	CAAGGGCGCG	1500
	AGAGCTGCTC	CTGCTGGGTG	TAGGCAGTGA	GGCCGTGCGG	GCAGAGCTGG	AGGAGATCAC	1560
	AGGCAGCCCA	AAGCATGTGA	TGGTCTACTC	GGATCCTCAG	GATCTGTTC	ACCAATCCC	1620
	TGAGCTGCAG	GGGAAGCTGT	GCAGCCGSCA	GCGGCCAGGG	TGCCGGACAC	AAGCCCTGGA	1680
20	CCTCGTCTTC	ATGTTGGACA	CCTCTGCCCTC	AGTAGGGCCC	GAGAATTTTG	CTCAGATGCA	1740
	GAGCTTTGAG	AGGAAGCTGTG	CCCTCCAGTT	TGAGGTGAAC	CCTGACGTGA	CACAGGTCCG	1800
	CCTGTGGTG	TATGGCAGCC	AGGTGCAGAC	TGCCTTGGG	CTGGACACCA	AACCCACCGG	1860
	GGCTGCGATG	CTGCGGGCCA	TTAGCCAGGC	CCCCACCTTA	GGTGGGCTGG	GCTCAGCCGG	1920
	CACCGCCCTG	CTGCACATCT	ATGACAAAGT	GATGACCGTC	CAGAGGGGTG	CCCCGCCCTG	1980
25	TGTCCCAAA	GCTGTGGTGG	TGCTCACAGG	CGGGAGAGGC	GCAGAGGATG	CAGCCGTTC	2040
	TGCCCAAGG	CTAGGAACA	ATGGCATCTC	TGTCTTGGTC	GTGGCGTGG	GGCCTGTCTC	2100
	AAGTGAGGT	CTGCGGAGGC	TTGCAGGTCC	CGGGATTTC	CTGATCCACG	TGGCAGCTTA	2160
	CGCGGACCTG	CGGTACCACC	AGGACGTGCT	CATTGAGTGG	CTGTGTGGAG	AAGCCAAAGCA	2220
	GCCAGTCAAC	CTCTGCAAA	CCAGCCCGTG	CATGAATGAG	GGCAGCTGGC	TCCTGCAGAA	2280
30	TGGGAGCTAC	CGCTGCAAGT	GTGCGGATGG	CTGGGAGGGC	CCCCACTGCG	AGAAACCGATT	2340
	CTGAGACCG	CCCTGAGGCA	CATGGCTCCC	GTGCAGGAGG	GCAGCAGCGG	TACCCCTCCC	2400
	AGCAACTACA	GAGAAGCCCT	GGGCACTGAA	ATGGTGCCTA	CCTTCTGGAA	TGTCTGTGCC	2460
	CCAGGTCCCT	AGAATGTCTG	CTTCCCGCGG	TGGCCAGGAC	CACATATTCTC	ACTGAGGGAG	2520
	GAGGATGTCC	CAACTGCAGC	CATGCTGCTT	AGAGACAAGA	AAGCAGCTGA	TGTCACCCAC	2580
35	AAAGCTGTTT	GTGAAAGT	TTTGTATGTT	AAGTAAATAC	CCACTTTCTG	TACCTGTCTG	2640
	GGCTTGTGTA	GGCTATGTCA	TCTGCCACCT	TTCCCTTGAG	GATAAACCAAG	GGGTCTCTGAA	2700
	GACTTAAATT	TAGCGGCGCTG	ACGTTCCCTT	GCACACATC	AATGCTCGCC	AGAATGTTGT	2760
	TGACACAGTA	ATGCCAGCA	GAGGCCTTTA	CTAGAGCATC	CTTTGACGG		

Seq ID NO: 445 Protein sequence
Protein Accession #: Eos sequence

45	1	11	21	31	41	51	
	MPPFLLEAV	CVFLFSRVPP	SLPLQEVHVS	KETIGKISAA	SKMMCSAAV	DIMPLLDGSN	60
	SVGKGSFERS	KHFAITVCDG	LDISPERVRV	GAFQPSSTPH	LEFPLDSFST	QQEVKARIKR	120
	MVFKGRTET	ELALKYLLHR	GLPGGRNASV	PQILIIIVTDG	KSQGDVALPS	KQLKERVTV	180
	FAVGVRFPFW	EELHALASEP	RQHVHLLAEQ	VEDATNGLFS	TLSSSAICSS	ATPDCRVEAH	240
	PCEHRTLEMV	REFAGNAPCW	RGRRTLAVL	AAHCPFYSWK	RVPLTHPATC	YRTTCPGPCD	300
50	SQPCQNGGT	VPEGLDGYQC	LCPLAFGGEA	NCAKLKSLC	RVDLLFLDLS	SAGTTLDGFL	360
	RAKVFVRKRV	RAVLSSEDSRA	RVGVATYSRE	LLVAVPVGEY	QDVDPDLVNSL	DGIPFRGGPT	420
	LTGSALRQAA	ERGFSGSATRT	QDRPRRVVV	LLTESHSEDE	VAGPARHARA	RELLLLGVGS	480
	EAVRAELEEI	TGSPKHMVMY	SDPDQLFNQI	PELQGLKCSR	QRPGCRTQAL	DLVFMLODSA	540
	SVGPNFAGM	QSFVRSCALQ	FEVNPVDVTQ	GLVVYGSQVQ	TAFGLDTPKT	RAAMLRAISQ	600
55	APYLGVGSA	GTALLHIYDK	VMTVQRGARP	GVPKAVVVL	GGRGAEADA	PAQLKRNNGI	660
	SVLVGVGV	LSEGLRRLAG	PRDSLHVA	YADLRYHQDV	LLENLCEAK	QPVNLCKPSP	720
	CMNEGSCVLIQ	NGSYRCKCRD	GWEGPHCENR	FLRRP			

Seq ID NO: 446 DNA sequence
Nucleic Acid Accession #: NM_031942.1
Coding sequence: 145..1260

65	1	11	21	31	41	51	
	CCCCAGCCCC	GCCCTCCCGG	GCCCGGGTCG	GCGCGCCCGAG	CCTGCCAGCC	GCGCTGCTGC	60
	TGCTCCTCCT	GCTGTGGGAC	CGCTGACCGC	GCGGCTGCTC	CGCTCTCCCC	GCTCCAAGCG	120
	CGATCTGGG	CACCGGCCAC	CAGCATGGAC	GCTCGCCGGG	TGCGCGAGAA	AGATCTCAGA	180
	GTAAGAAGA	ACTTAAAGAA	ATTGAGATAT	GTGAAGTTGA	TTTCCATGGA	AACCTCGTCA	240
	TCCTCTGATG	ACAGTTGTGA	CAGCTTTGCT	TCTGATAATT	TTGCAAAAC	GAGGCTGCAG	300
	TCAGTTCGGG	AAGCGTGTAG	GACCCGCGAGC	CAGTGCAGGC	ACTCTGGACC	TCTCAGGGTG	360
70	GCGATGAAGT	TTCCAGCGCG	GAGTACCAGG	GGAGCAACCA	ACAAAAAAGC	AGAGTCCCGC	420
	CAGCCCTCAG	AGAATTCTGT	GACTGATTCC	AACCTCGATT	CAGAAGATGA	AAGTGAATG	480
	AATTTTTTGG	AGAAAAGGCG	TTTAAATATA	AAGCAAAACA	AAGCAATGCT	TGCAAAACTC	540
	ATGTCTGAAT	TAGAAAGCTT	CCCTGGCTCG	TTCCGTGGAA	GACATCCCC	CCCAGGCTCC	600
	GACTCACAAT	CAAGGAGACC	GCGAAGGCGT	ACATTCCCGG	GTGTTGCTTC	CAGGAGAAAC	660
75	CCTGAACGGA	GAGCTCGTCC	TCTTACCAGG	TCAAGGTCCC	GGATCCTCGG	GTCCTCTGAC	720
	GCTCTACCCA	TGGAGGAGGA	GGAGGAAGAG	GATAAGTACA	TGTTGGTGAG	AAAGAGGAAG	780
	ACGTTGATG	GCTACATGAA	TGAAGATGAC	CTGCCCGAG	GCGTCCGCTC	CAGATCATCC	840
	GTGACCCCTT	CGCATATAT	TGCCCCAGTG	GAAGAAATTA	CAGAGGAGGA	GTTGGAGAAC	900
	GTCTGCAGCA	ATTCTCGAGA	GAAGATATAT	AACCGTTTAC	TGGGCTCTAC	TGTGTCATCA	960
80	TGCCGTGAGA	AGACTATTGA	TACCAAAACA	AACCTGCAGAA	ACCCAGACTG	CTGGGGCGTT	1020
	CGAGGCCAGT	TCTGTGGCCC	CTGCCTTCGA	AACCGTTATG	GTGAAGAGGT	CAGGGATGCT	1080
	CTGCTGGATG	CGAACTGGCA	TTGCCCGCCT	TGTCGAGGAA	TCTGCAACTG	CAGTTTCTGC	1140
	CGGCAGCGAG	ATGACGCGTG	TGCGACTGGG	GTCCCTGTGT	ATTTAGCCAA	ATATCATGGC	1200
	TTTGGGAATG	TGACTGCCTA	CTGAAAAGC	CTGAAAAGC	AATTTGAAAT	GCAAGCATAA	1260
85	TATCTGAAAA	ATTGTGCTGC	TGCCTTCTAC	TTCTCAAAATC	TTTCTGTGAA	AAGTTTCCAA	1320
	TTTCTTCACT	GAAACCTGAG	TTAAAAATCT	TGATGATCAG	CCTGTTTCAT	AGAATACTCC	1380
	AATCAAGTTA	ATCTTAGCAG	ACATGTGTTT	CTGGAGCATC	ACAGAAGGTA	TATTGCTAGT	1440

TACACTTGGC CCTCTGTCAG TTTCTTCTCT GCTCCCAACC OCCATCTCAT AGCATCCCC 1500
 TCTATTTCCA ATGCTCTCTT CCAACCGCTT AGTTTCTGAA TTTCTTTTAA ATTACAGTTT 1560
 TATGAAAGCA TATTTTATTT ACTTGGTGTT GAAATAGCCC TCATAAAACC TAAGCAGTTG 1620
 GAAACACAAAT AATAGTATTA ACTAACTAGA TCTATTGAAT TTCAGAGAAG AGCCTTCTAA 1680
 CTGTGTTTACA CAAAACGAG: TATGATTTAG CACTCATACT AGTTGAAATT TTTAATAGAA 1740
 TCAAGGCACA AAAGTCTTAA: AACCATGTGG AAAAATTAGG TAATTATTGC AGATTGATGT 1800
 CTCTCAATCC CATGTATTGC GCTTATGITA CAAGTTGTTG TCACAGTTGA GACTTAATTT 1860
 CTCTTAATTT CTCTGCGCCG AAGGGTAAGT GGTGCGTCCA GCTTACACGA TCATAATTCA 1920
 AAGGTTGAGT GGCATGTAA: TACTTAATTA AAATAATGAT GGAAGAGCTA TCTGGAGATT 1980
 ATGAGTAAAG: TGATTGTAAT TTTCAATATA AAACCTTAGT ATAATTGTAG TTTGCAAGT 2040
 TTATTTTCACT TCACATGTAA GGTATTGCAA ATAAATTCCT GGCAATTTT GTATGGAAC 2100
 TTGATATTAA AAAGTATGCT GTGGTTCTTT CGAGTTTCTT GTAAATTTAT AAACAGGCA 2160
 CAAGGTTTCAA GTTTAGATT TTAAGCATT: TATAACATG ATAAGTGCTT TTTTGGAGAT 2220
 GTAACCTTAA GCACTTTGTT AACCTGACAT CTCTGCCAGT CTAGTTTCTG GGCAGGTTT 2280
 CTGTGTCAGT ATTCCCCCTC CTCTTTGCAT TAATCAAGGT ATTTGGTAGA GGTGGAATCT 2340
 AAGTGTGTTG ATGTCCAATT TACTTGCATA TGTAAACCAT TGCTGTGCCA TTCAATGTTT 2400
 GATGCATAAT TGGACCTTGA ATCGATAAGT GTAAATACAG CTTTGTATCT GTAATGCTTT 2460
 TATACAAAG: TTTATTTTAA TAATAAAATG TTTGTTCTAA AAAAAAATA

Seq ID NO: 447 Protein sequence
 Protein Accession #: NP_114148.1

1 11 21 31 41 51
 MDARRVPQKD LRVKLNKKF RVVKLISMET SSSSDSDSDS FASDNFANTR LQSVREGCRT 60
 RSQCRHSGLP RVAMFPPARS TRGATNKKAE SRQPSNSVT DSNDSDEDES GNFLEKRAL 120
 NIKQNKAMLA KLMSLESEFP GSFRGRHPLP GSDSQSRPR RRTFPGVASR RNPERRARPL 180
 TRSRRLILGS LDALPMEEEE EEDKYMLVRK RKTVDGYMNE DDLPRSRRSR SSVTLPHIIR 240
 PVBEITEEL EIVCSNSREK IYNRLSGSTC HQCRQKTDIT KTNCRNPDCW GVRGQFCGPC 300
 LRNRYGESEVR DALLDPNWHC PPCRGICNCS FCRQRDGRCA TGVLYLAKY HGFNVHAYL 360
 KSLKQEFEMQ A

Seq ID NO: 448 DNA sequence
 Nucleic Acid Accession #: NM_019894
 Coding sequence: 1..1314

1 11 21 31 41 51
 ATGTTACAGG ATCCTGACAG TGATCAACCT CTGAACAGCC TCGATGTCAA ACCCCTGGGC 60
 AAACCCCGTA TCCCCTAGGA GACCTTCAGA AAGGTGGGGA TCCCATCAT CATAGCACTA 120
 CTGAGCCTGG CGAGTATCAT CATTTGTGTT GTCCATCA AGGTGATTCT GGATAAATAC 180
 TACTTCTCTT GCGGCGAGCC TCTCCACTTC ATCCCGAGGA AGCAGCTGTG TGACGGAGAG 240
 CTGGAGTCTC CTTGGGGGGA GGACGAGGAG CACTGTGTCA AGAGCTTCCC CGAAGGGCCT 300
 GCAGTGGCAG TCGGCTCTC: CAAGGACCGA: TCCACACTGC AGGTGCTGGA TCGGCCACA: 360
 GGGAACTGGT TCTCTGCTGT TTTGACAAC TTCACAGAAG CTCTCGCTGA GACAGCCTGT 420
 AGGCAGATGG: GCTACAGCAG: CAAACCCACT TTCAGAGCTG TGGAGATTGG CCCAGACCAG 480
 GATCTGGATG TTTTGTAAAT CACAGAAAAC AGCCAGGAGC TTGCGATGCG GAACCTCAAGT 540
 GGGCCCTGTC TCTCAGGCTC CTTGCTCTCC CTGCACTGTC TTGCTGTGG GAAGAGCCTG 600
 AAGACCCCCC GTGTGTGGGG TGGGGAGGAG GCCTCTGTGG ATTCTTGCCC TTGGCAGGTC 660
 AGCATCCAGT: AGCAACAACA: GCACGTCTGT GGAGGGAGCA TCCCTGGACC CCATCGGCTC 720
 CTCACGCGAG CCCACTGCTT CAGGAAACAT ACCGATGTGT TCAACTGGAA GGTGGGGCA 780
 GGCTCAGACA AACTGGGCGC CTTCCCATCC CTGGCTGTGG CCAAGATCAT CATCAATTGA 840
 TTCAACCCCA TGTACCCCAA AGACAATGAC ATCGCCCTCA TGAAGCTGCA GTTCCCACTC 900
 ACTTTCTCAG GCACAGTCAG GCCCATCTGT CTGCCCTTCT TTGATGAGGA GCTCACTCCA 960
 GCCACCCGAC TCTGATCAT TGGATGGGGC TTTACGAAGC AGAATGGAGG GAAGATGTCT 1020
 GACATACTGC TCAGGGGTC AGTCCAGGTC ATTGACAGCA CACGGTGCAA TGCAGACGAT 1080
 GCGTACCAGG GGGAGAGTCA CGAGAAGATG ATGTGTGCAG GCATCCCGGA AGGGGGTGTG 1140
 GACACCTGCC AGGGTGACAG TGGTGGGCCC CTGATGTACC AATCTGACCA GTGGCATGTG 1200
 GTGGGCTGCT TTAGCTGGGG CTATGGCTGC GGGGGCCCGA GCACCCAGG AGTATACACC 1260
 AAGGTCTCAG CCTATCTCAA CTGGATCTAC AATGTCTGGA AGGCTGAGCT GTAA

Seq ID NO: 449 Protein sequence
 Protein Accession #: NP_063947.1

1 11 21 31 41 51
 MLQDPDSQDP LNSLDVKPLR KPRIPMETFR KVGIPPIIAL LSLASIIIV VLIKVILDKY 60
 YFLCGQLHP IPRQLCDGE LDCPLGEDEE HCVKSPPEGP AVAVRLSKDR STLQVLDSAT 120
 GNWFSACFDN FTEALAEAC RQMGYSKPT FRAVEIGPDQ DLDVVEITEN SQELMRNNS 180
 GPCLSGSLVS LHCLACGKSL KTRPVVGEE ASVDSWPNQV SIQYDKQHVC GGSILDPHVV 240
 LTAHCFRKH TDVFNWVRA GSDKLGSPFS LAVAKIIIE FNPMPKDND IALMKLQFPL 300
 TFSGTVRPIC LPPFDEELTP ATPLWIIWGM FTKQNGGKMS DILLQASVQV IDSTRCNADD 360
 AYQGEVTEKM MCAGIPEGGV DTCQDGGGP LMYQSDQHWV VGIVSWGYGC GGPSTPGVYT 420

Seq ID NO: 450 DNA sequence
 Nucleic Acid Accession #: XM_051860.2
 Coding sequence: 52..3042

1 11 21 31 41 51
 GCTCACCAG GAAAAATATG CAATGCTCCC ATTGATATAC AGGCCACTAC AATGGATGGA 60
 GTTAACTCA GCACCGAGGT TGTCTACAAA AAAGGCCAGG ATTATAGGTT TGCTTGTCTAC 120
 GACCGGGCA GAGCTGCCG GAGCTACCGT GTACGGTTCC TCTGTGGGAA GCCTGTGAGG 180
 CCCAACTCA CAGTACCAT TGACACCAAT GTGAACAGCA CCATTCTGAA CTGGAGGAT 240
 AATGTACAGT CATGAAAACC TGGAGATACC CTGCTCATG CCAGTACTGA TTACTCCATG 300
 TACCAGGCA: AAGAGTTCCA: GGTGCTTCCC TGCAGATCCT GGGCCCCCAA CCAGGTCAAA 360

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

GTGGCAGGGA: AACCAATGTA CTTGCACATC GGGGAGGAGA TAGACGGCGT GGACATGCGG 420.
GGGAGGTTG GGCCTTCTGAG: CCGGAACATC ATAGTGATGG GGGAGATGGA GGACAAATGC 480.
TACCCCTACA GAAACCATCAT CTGCAATTTT TTTGACTTGG ATACCTTTGG GGGCCACATC 540.
AAGTTTGTCT TGGGATTTAA GGCAGACAC TTTGAGGGCA CGGAGCTGAA GCATATGGGA 600.
CAGCAGCTGG TGGGTTCAGTA CCGATTCAC TTCCACTGG CCGGTGATGT AGACGAAAGG 660.
GGAGGTTATG: ACCCAACCAT ATACATCAGG GACCTCTCCA TCCATCATAC ATTCTCTCGC 720.
TGCTGCACAG TCCGATGGCTC CAATGGCTTG TTGATCAAGG AGTTTGTGGG CTATAACTCT 780.
TTGGGCCACT GCTTCTTAC: GGAAGATGGG CCGGAGGAAC GCAACACTTT TGACCACTGT 840.
CTTGGCTTCC TTGTCAAGTC TGSAAACCTC CTCCCTCGG ACGTGACAG CAAGATGTGC 900.
AAGATGATCA: CAGGAGACTC CTACCCAGGG TACATCCCA AGCCAGGCA AGACTGCAAT 960.
GCTGTGTCCA CCTTCTGGAT GGCCAATCCC AACACAACC TCATCAACTG TGCCGCTGCA 1020.
GGATCTGAGG AAACCTGATT TTGGTTTATT TTTCACCAG TACCAACGGG CCCCTCOGTG 1080.
GGAATGTACT: CCCCAGGTTA: TTCAGAGCAC ATTCCACTGG GAAAATTCTA TAACAACCGA 1140.
GCACATTCCA ACTACCGGC: TGGCATGATC ATAGACAAGG GAGTCAAAAC CACCGAGGCC 1200.
TCTGCCAAGC: ACAAGCGGCG GTTCTCTCA: ATCATCTCTG CCAGATACAG: CCTCACCAG 1260.
GACGCCGACC CCGTGAAGCC CCGGAGGCGG GCCATCATCA GACACTTCAT: TGCCTACAAG 1320.
AACGAGACC: ACGGGGCTG: GCTGCGCGG: GGGGATGTGT GGCTGGACAG: CTGCGGTTT 1380.
GCTGACATG GCATTGGCCT GACCTGCGCC AGTGGTGGAA: CCTTCCCGTA: TGACGACGCG 1440.
TCCAAGCAAG AGATAAGAA: CAGCTTGTCT GTTGGCGAGA: GTGGCAACGT GGGGACGGAA 1500.
ATGATGAGAG: CAGGATCTG: GGGCCTGCG: GGCTTGGACC ATAGCGGAAG: GACCTCCCT 1560.
ATAGGCGAGA: ATTTTCCAAT: TAGAGGAATT: CAGTTATATG ATGGCCCCAT CAACATCCAA 1620.
AATGCACTT: TCCGAAGATT: TGTGGCCTG: GAGGCGCGG: ACACAGCGG: CTGGCCTTC 1680.
CGCCTGAATA ATGCTTGGCA: GAGCTGCCCC CATAACAAG TGACCGCAT TGCTTTGAG 1740.
GACGTTCCGA TTACTTCCAG AGTGTCTTC: GGGAGGCTG: GGCCCTGGTT CAACAGCTG 1800.
GACATGAGAG: GGGATAAGAC: ATCTGTGTT: CATGACGTG: ACGCTCCGT GTCCGAGTAC 1860.
CCTGGCTCT: ACCTCAGGAA: GAATGACAAC TGGCTGGTCC GGCACCCAGA: CTGCATCAAT: 1920.
GTTCCGACT: GAGAGGGG: CATTTGCAGT: GGGTGTATG CACAGATGTA CATTCAAGCC 1980.
TACAAGACCA GTAACCTGCG: AATGAAGATC ATCAAGAATG ACTTCCCGAG CCACCTCTT 2040.
TACCTGAGAG: GGGCGCTCAC: CAGGAGCACC CATTACCAGC AATACCAACC GGTGTGACC 2100.
CTGCAAGAG: GCTACACCAT: CCACTGGGAC CAGACGGGCC CGCCGAACT: GGCATCTGG: 2160.
CTCATCACT: TCAACAAGG: CGACTGGATC CGAGTGGGCG: TCTGTACCC GCGAGGCACC 2220.
ACATTTCCA: TCTCTCGGA: TGTTCACAAT CGCTGCTGA: AGCAACGCT CAAGACGGG: 2280.
GTCTTCTGA: GGACCTTGA: GATGGACAAA GTGGAGCAGA: GCTACCTGG CAGGAGCCAC: 2340.
TACTACTGGG ACGAGGACT: AGGGCTGTG: TTCTGAAGC TGAAGCTCA GAACGAGAGA: 2400.
GAGAAGTTG: CTTTCTGCTC: CATGAAAGGC TGTGAGGGA: TAAAGATTAA: AGCTCTGATT: 2460.
CCAAAGAACG CAGCGCTCAG: TGAATGACA: GCCACAGCTT ACCCAAGTT: CACCGAGAG: 2520.
GCTGTCTAG: ACGTCCGAT: GCGCAAGAAG CTCTTTGGTT: CTCAGCTGAA: AACAAAGGAC: 2580.
CATTTCTTG AGGTGAAGAT GGAGAGTTCC AAGCAGCACT TCTTCCACT CTGGAACGAC: 2640.
TTGCTTACA: TTGAAGTGA TGGGAAGAAG: TACCCAGTT: CGGAGGATGG CATCCAGGTG: 2700.
GTGGTGATTG ACGGGAACCA: AGGGCGCGTG GTGAGCCACA: CGAGCTTCA GAACTCCATT 2760.
CTGCAAGGCA: TACCATGGCA: GCTTTCAAC TATGTGGCGA CCATCCCTGA CAATTCATA 2820.
GTGCTTATGG: CATCAAAGG: AAGATACGTC TCCAGAGGCC CATGAGCAG AGTGCTGGA: 2880.
AAGCTTGGG: CAGACAGGG: TCTCAAGTTG: AAAGAGCAAA TGGCATTCGT TGGCTTCAA 2940.
GGCAGCTTCC GGCCTACTG: GGTGACACTG GACACTGAGG ATCACAAGC CAAATCTTC 3000.
CAAGTTGTG CCACTCCCTG: GGTGAAGAAG AAGAAGTTGT GAGGACAGCT GCCCGCCGT 3060.
GCCACCTGT: GGTAGACTAT GACGCTGACT: CTGGCAGCA: GACCAAGTGG GATGGCTGG 3120.
GTCCCGCAGC CCGTCCGAGC AGCTGCTGG: GAAGCGCGTG TTTCAAGCCT: GATGGGCAA 3180.
GGGAAGGCTA: TCAGAGACCC TGGTCTGCC ACCTGCCCTT ACTCAAGTGT CTACCTGGAG: 3240.
CCCCCTGGG: GGTGCTGGCC AATGCTGGAA ACATTCACCT TCCTGAGCC TCTTGGGTG: 3300.
TTCTCTCTTA: TTTGCTGCTT TTAGTGGG: GTTGGGGGAC CATATCAGGA GACCTGGGTT 3360.
GTGCTGACAG: CAAAGATCCA: CTTTGGCAGG: AGCCCTGACC CAGCTAGGAG GTAGCTGGA 3420.
GGGCTGCTCA: TTCACAGAT: CCCATGTTCT TCAGCAGACA AGTGAGGGTG GTAAATGTAG 3480.
GAGAAAGAGC CTTGGCTTAA: AGGAAATCTT TACTCTGTGA: AGCAAGAGCC AACCTCACAG: 3540.
GATTAGGAGC TGGGGTAGAA CTGGCTATCC TTTGGGAAGA GGCAAGCCCT GCCTCTGGCC 3600.
GTGTCCACCT: TTCAGAGAG: TTTGAGTGGC AGGTTTGGAC TTGACTAGA TGACTCTCAA 3660.
AGGCCCTTTT AGTTCTGAGA: TTTCAAGAAAT CTGTGCAAT TCACATGGTA CCTGGAACCC 3720.
AACAGTTTCA: GGAATCCAT: TGATATCCAT: GATGCTGGGT: GCCCAGCGC ACACGGGATG 3780.
GAGAGGTGAG AACTAATGCC TAGCTTGAGG GCTCTGCACT: CAGTAGGGC AGGCAGTCAG: 3840.
GTCCATGTGC ACTGCAATGC CAGGTGGAGA AATCACAGAG AGGTAAATG GAGGCCAGTG: 3900.
CCATTTCA: GGGGAGGCTC: AGGAAGGCTT CTTGCTTACA: GGAATGAAGG CTGGGGCAT 3960.
TTTGTGGGG: GGAGATGAGG CAGCCTCTGG AATGGCTCAG: GGAATCAGCC CTCCTGCGG 4020.
CTGCTGCTG: AAGCTGTGTA: CTACGGGTC: GCCCTTTGCT CACGTCTCT TGGCCACTC 4080.
ATGATGGAGA AGTGTGTGTA GAGGGGAGCA ATGGGCTTTG CTGCTTATGA GCACAGAGGA: 4140.
ATTCAGTCCC CAGGAGCCCC TGCTCTGAC TCCAAGAGG: TGAAGTCCAC AGAAGTGAGC 4200.
TCTGCTTGA: GGGCTCAT: TGCTCTCAT CCAGGGAAC: GAGCAGAGG GGCCTCCAGG 4260.
AGACCTAGA: TGTGCTGTA: CTCCTCGGC CTGGGATTTT: AGAGCTGGA ATATAGAAA 4320.
TATCTAGCCC AAAGCCTTCA: TTTTAAACAGA: TGGGAAAGT: GAGCCCCAA GATGGGAAG 4380.
AACACACAG CTAAGGGAGG: GCCTGGGGAG CCCCACCCTA: GCCCTTGTG CCACACCACA 4440.
TTGCTCAAC AACCGGCCCC AGAGTGCCCA: GGCACCTCTG AGGTAGCTTC TGGAAATGGG 4500.
GACAAGTCCC CTCGAAGGAA AGGAAATGAC TAGAGTAGAA TGACAGCTAG CAGATCTCTT 4560.
CCCTCTGCT: CCCAGGCA: ACAAACCGC: CCTCCCTTG GTGTGGCGG TCCCTGTGGC 4620.
CTTCACTTTG TTTCACTACT: GTCAGCCAG: CCTGGGTGCA CAGTAGCTGC AACTCCCAT 4680.
TGGTGCTACC TGGCTCTCT: GTCTCTGAG: CTCTACAGGT GAGGCCAGC AGAGGGAGTA 4740.
GGGCTGCGCA TGTCTCTGGT GAGCCAATTT GGCTGATCTT GGGTGTCTGA ACAGCTATTG 4800.
GGTCCACCCC AGTCCCTTTT: AGCTGCTGCT: TAATGGCCCTG CTCTCCCTT GGGCCACTT 4860.
ATAGAGAGCC CAAAGAGCTC CTGTAAGAGG: GAGAATCTA TCTGTGGTTT: ATAATCTTGC 4920.
ACGAGGCACC AGAGTCTCCC TGGGTCTGT: GATGAACATC: ATTTATCCCC TTTCTGCCC 4980.
CAACCAAAA CTCTTCTCT: CAAAGAGGGC: CTGCTGGCT: CCCCACCCC AACTGCACCC 5040.
ATGAGACTCG GTCTCAAGAT: CCAATCCCCA: GGTGGGAGCC AACTGTGAGG GAGGTCTTTC 5100.
CCACCAAAA CTCTTCAAGT: GCTGGGAGG: GACCATAGGG: CTCTGCTTTT: AAAGATATGG 5160.
CTGCTTCAA: GGCCAGATC: ACAGGAAGGA: CTTCTTCCAG GGAGATTAGT GGTGATGGAG 5220.
AGGAGAGTTA: AATGACCTC: ATGCTCTCT: TGTCCACGCT TTTGTGAGT TTTCACTCTT 5280.
CTAATGCAAG GGTCTCACAC: TGTGAACCC: TTAGGATGTG ATCACTTTCA GGTGGCCAGG 5340.
AATGTTGAAT GTCTTTGGCT: CAGTTCAATT: AAAAAAGATA TCTATTGAA AGTTCTCAGA 5400.
GTGTACATA TGTCTCACAG: TACAGGATCT: GTACATAAAA: GTTCTTTCC TAAACCATTC 5460.
ACCAAGAGCC: AATATCTAGG: CATTTCTTG: GTAGCAGAAA: TTTCTTATT: GCTTAGAAA 5520.
TTGTCTCTCT: TGTATTCTG: GTTGTGAAGA: CTTAAGTGAG: TTAGGCTTTT: AAGGAAAGCA 5580.

ACGCTCTCT GAAATGCTTG TCITTTTCT GTTGCAGAAA TAGCTGGTCC TTTTTCGGGA 5640
 GTTAGATGTA TAGAGTGTTT GTATGTAAC ATTCTTGTA GGCATCACC TGAACAAAGA 5700
 TATATTTTCT ATTTATTTAT TATATGTGCA CTTCAGAAAG TCACTGTCTAG AGAAATAAAG 5760
 AATTGCTTA AATGTCAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 451 Protein sequence
 Protein Accession #: XP_051860.2

1 11 21 31 41 51
 MDGVNLSTEV VYKKGQDYRF ACYDRGRACR SYRVRFCLGK FVRPKLTVTI DTNVNSTILN 60
 LEDNVQSNKP GDTLVIASDT YSMYQAEFQ VLPSCRSCAPN QVKVAGKPMY LHIGEEIDGV 120
 DMRAEVGLLS RNIIIVMGEME DKCYPYRNHI CNFFDFDTFG GHKIFALGFK AAHLEGTELK 180
 HMQQLVGGY PIHFLHAGDV DERGGYDPT YIRDLSEIHT FSRCTVHGS NGLLIKDVVG 240
 YNSLGHCFPT EDGPEERNTF DHCLGLLVKS GTLLPSDRDS KMKCMITGDS YPGYIKPKRQ 300
 DCNAVSTFWM ANPNVNLINC AAAGSEETGF WFIHHEVPTG PSVGMYSPTY SEHIPLGKPY 360
 NNRAHSNYRA GMIDNGVKT TEASAKDKRP FLISISARYS PQQADAPLKP REPAIIRHFI 420
 AYKNDHGW LRGDVLWLS CRPADNGIGL TLASGGTFY DDGSKQEIKN SLFVGSSENV 480
 GTEMMDNRW GPGLDHSGR TLPIGNFPI RGIQLYDGP NIQNTCTFRK VALEGRHTSA 540
 LAFRLNNAWQ SCFNNVTGI AFEDVPITSR VFFGEPGPMF NQLDMGDKT SVPHVDVGSV 600
 SEYPGSYLTK NDNMLVRHPD CINVPDWRGA ICSGCAQMY IQAYKTSNLR MKIKNDFPS 660
 HPLYLEGALT RSTHYQQYQP VVTLQKGYTI HWDQTAPEL AIWLINFNGK DWIRVGLCPY 720
 RGTTFSLSD VNRLLKQTS KTGTVFRTLQ MDKVEQSYFG RSHYYWDEDS GLFLKLLKAQ 780
 NEREKFAFCS MKGCEKIKK ALIPKNAGVS DCTATAPKF TERAVVDVPM PKLFGSQLK 840
 TKDHFLEVKM ESSKHFFHL WNDFAIYEDV GKYPSSSDG IQVVVIDGNQ GRVVSHTSFR 900
 NSILQGIPOW LFNYVATIPD NSIVLMASKG RYVSRGPWTR VLEKLGADRG LKLKQMAFV 960
 GFKGSFRPIW VTLDTEHKA KIFQVVPFV VKKKKL

Seq ID NO: 452 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 261..2861

1 11 21 31 41 51
 GAGCTAGCGC TCAAGCAGAG CCCAGCGCGG TGCTATCGGA CAGAGCCTGG CGAGCGCAAG 60
 CGGCGCGGG AGCCAGCGGG GCTGAGCGCG GCCAGGGTCT GAACCCAGAT TTCCAGACT 120
 AGCTACCACT CGCTGTGCC ACGCCCGCGG AGCTCGCGGC GCTCGCGGT CAGCGACACC 180
 ACGTCCGGG CGCTGCGCT CCGTGGCGGC GAGGCGTGAC ACTGTCTCGG CTACAGACCC 240
 AGAGGAGCA CACTGCCAGG ATGGGAGCTG CTGGGAGGCA GGACTTCCTC TTCAAGGCCA 300
 TGCTGACCAT CAGCTGGCTC ACTCTGACCT GCTTCCCTGG GGCACATCC ACATGGCTG 360
 CTGGGTGCCG TGACAGAGC CCGAGTTGC AACCTGGAA CCTGGCCAT GACCAAGACC 420
 ACCATGTGCA TATCGGCCAG GGCAAGACAC TGCTGCTCAC CTCTTCTGCC ACGGTCTATT 480
 CCATCCACAT CTCAGAGGGA GGCAAGCTGG TCATTAAAGA CACGAGCAG CCGATTGTTT 540
 TCGCAACCGG GCACATCCTG ATTGACAACG GAGGAGAGCT GCATGCTGGG AGTGCCCTCT 600
 GCGCTTTCCA GGGCAATTC ACCATCATT TGTATGSAAG GCGTATGAA GGTATTACAG 660
 CGGATCTCTA CTATGGTCTG AGGTACATTG GGGTTGGTAA AGGAGGCGCT CTTGAGTTGC 720
 ATGGAACGAA AAGACTCTCC TGGACATTTC TGAACAAGAC CCTTCAACCA GGTGGCATGG 780
 CAGAAAGGAG CTATTCTTTT GAAAGGAGCT GGGGCCACCG TGGAGTTATT GTTCATGTCA 840
 TCGACCCCA ATCAGGCACA GTCATCCATT CTGACCGGTT TGACACCTAT AGATCCAAGA 900
 AAGAGAGTGA ACGTCTGGTC CAGTATTGTA ACGCGGTGCC CGATGGCAGG ATCCTTTCTG 960
 TTGCAGTGAA TGATGAAGGT TCTCGAAATC TGGATGACAT GGCCAGGAAG GCGATGACCA 1020
 AATTGGGAAG CAACACCTTC CTGACCTCTG GATTTAGACA CCTTGGAGT TTTCTAAGCT 1080
 TGAAGGAAAT TCCATCATCT TCAGTGGAA AGCATATTGA ATATCATGGA CATCGAGGCT 1140
 CTGCTGCTGC CCGGTATTC AAATTGTTC AGACAGAGCA TGGGAATAT TTCAATGTTT 1200
 CTTTGTCCAG TGATGGGTT CAAGACGTGG AGTGGACGGA GTGGTTCGAT CATGATAAAG 1260
 TATCTCAGAC TAAAGTGGG GAGAAAATTT CAGACCTCTG GAAAGCTCAC CCAGGAAAAA 1320
 TATGCAATCG TCCCATTTGAT ATACAGGCCA CTACAATGGA TGGAGTTAAC CTCAGACCG 1380
 AGGTTGTCTA CAAAAAGCG CAGGATTATA GGTTTGCTTG TCAACAGCGG GCGAGAGCCT 1440
 GCCGAGCTA CCGTGTACGG TTCTCTGTG GGAAGCTGT GAGGCCCAAA CTCACAGTCA 1500
 CCATTGACAC CAATGTGAAC AGCACCATTC TGAACCTGGA GGATAATGTA CAGTCATGGA 1560
 AACCTGAGGA TACCTGGTC ATTGGCAGTA CTGATTACTC CATGTACCAG GCAGAAAGAT 1620
 TCCAGGTGCT TCCCTGCAGA TCCCTGCCCC CCAACAGGTT CAAAGTGGCA GGGAAACCAA 1680
 TGTACCTGCA CATCGGGAG GAGATAGACG GCGTGGACAT GCGGGCGGAG GTTGGGCTTC 1740
 TGAGCGGAA CATCATAGTG ATGGGGGAGA TGGAGGACAA ATGCTACCCC TACAGAAACC 1800
 ACATCTGCAA TTTCTTTTAC TTGATACCT TTGGGGGCCA CATCAAGTTT GCTCTGGGAT 1860
 TTAAGGCAGC ACACCTGGAG GGCACGGAGC TGAAGCATAT GGGACAGCAG CTGGTGGGTC 1920
 AGTACCCGAT TCACCTCCAC CTGGCCGGTG ATGTAGACGA AAGGGGAGGT TATGACCCAC 1980
 CCACATACAT CAGGGACCTC TCCATCCATC ATACATTCTC TCGCTGCGTC ACAGTCCATG 2040
 GCTCCAATG CTGTGTGATC AAGGACGTTG TGGGCTATAA CTCTTTGGGC CACTGCTTCT 2100
 TCACGGAAGA TGGCCGGAGG SAACGCAACA CTTTGAACA CTGTCTGGC CTCTTGTCA 2160
 AGTCTGGAAC CTTCTCCCC TCGGACCGTG ACAGCAAGAT GTGCAAGATG ATCACAGAGG 2220
 ACTCCTACCC AGGGTACATC CCCAAGCCCA GGCAGACTG CAATGCTGTG TCCACCTTCT 2280
 GGATGGCCAA TCCCAACAAC AACCTCATCA ACTGTGCCGC TGCGAGATCT GAGGAAACTG 2340
 GATTTTGGIT TATTTTTCAC CAGTACCAA CGGGCCCCCT CGTGGGAATG TACTCCCCAG 2400
 GTTATTGAGA GCACATTCCA CTGGGAAAT TCTATAACA CGAGCACAT TCCAATACC 2460
 GGGCTGCGAT GATCATAGAC AACGAGTCA AAACCAACGA GGCCTCTGCC AAGGACAAGC 2520
 GGGCGTTCCT CTCATCATC TCTGCCAGAT ACAGCCCTCA CCAGGAGGCC GACCCGCTGA 2580
 AGCCCGGGA GCCCGCATC ATCAGACACT TCATTGCGTA CAAGAACCAG GACCAAGGGG 2640
 CCTGGCTGCG CCGCGGGGAT GTGTGGCTGG ACAGCTGCCA TTTCAAGGGG GAGGCTCAGG 2700
 AAGGCTCTTT GCTTACAGGA ATGAAGGCTG GGGCAITTT GCTGGGGGA GATGAGGCAG 2760
 CCTCTGGAAT GGCTCAGGGA TTCAGCCCTC CTGCGCGTGG CCGTCTGAAG CTGGTACTA 2820
 CGGGGTGCGC CTTTGTCTAC GTCTCTCTGG CCCACTCATG ATGAGAGAAG GTGGTCAGAG 2880
 GGGAGCAATG GCGTTTGTCT CTTATGAGCA CAGAGGAATT CAGTCCCCAG GCAGCCCTGC 2940
 CTCTGACTCC AAGAGGTGA AGTCCACAGA AGTGAGCTCC TGCCCTAGGG CCTCATTTGC 3000
 TCTTATCCA GGGAACTGAG CACAGGGGGC CTCCAGGAGA CCTAGATGT GCTGTACTC 3060
 CCTCGGCTG GATTTTCAGA GCTGGAATA TAGAAAATAT CTAGCCCAAA GCCTTCATTT 3120

5
10
15
20
25

```

TAACAGATGG GGAAGTGAG CCCCAGAT GSGAAGAAC CACACAGCTA AGGGAGGGCC 3180
TGGGAGAGCC CACCCTAGCC CTGCTGCCA CACCACATGG CCTCAACAAC CGGCCCCAGA 3240
GTGCCAGGCG ACTCCTGAGG TAGCTTCTGG AAATGGGGAC AAGTCCCCCTC GAAGGAAAGG 3300
AAATGACTAG AGTAGAATGA CAGCTAGCAG ATCTCTTCCC TCCTGCTCCC AGCGCACACA 3360
AACCCGCTCT CCCCTGGGTG TTGGCGGTCC CTGTGGCCTT CACTTTGTTC ACTACCTGTC 3420
AGCCGAGCCT GGGTGCACTG TAGCTGCAAC TCCCATTGGT TGCTACCTGG CTCTCCTGTC 3480
TCTGCAGCTC TACAGGTGAG GCCCAGCAGA GGGAGTAGGG CTGGCCATGT TTCTGTGTAG 3540
CCAATTGGCG TGATCTTGGG TGTCTGAACA GCTATTGGGT CCACCCCACT CCCTTTTCAGC 3600
TGCTGCTTAA TGCCCTGTCT TCTCCCTGGC CCACCTTATA GAGAGCCCAA AGAGCTCCTG 3660
TAAGAGGGAG AACTCTATCT GTGGTTTATA ATCTTGCAAG AGGCACCAGA GTCTCCCTGG 3720
GTCTTGATGAT GAATACATTT TATCCCTTTT CCTGCCCAA CACAAACTC TTTCTTCAA 3780
AGAGGGCCTG CCTGCTCTCC TCCACCCAAC TGCACCCATG AGACTCGGTC CAAGAGTCCA 3840
TTCCCCAGGT GGGAGCCCAAC TGTGAGGAG GTCTTTCCCA CCAACATCTT TTCAGCTGCT 3900
GGGAGGTGAC CATAGGGCTC TGCTTTTAAA GATATGGCTG CTTCAAAGGC CAGAGTCACA 3960
GGAGGACTTT CTTCAGGGA GATTAGTGGT GATGGAGAGG AGAGTTAAAA TGACCTCATG 4020
TCCTTCTTGT CCACTGTTT GTTGAGTTT CACTTCTTA ATGCAAGGGT CTCACACTGT 4080
GAACCACTTA GGATGTGATC ACTTTGAGT GGCACAGAA GTTGAATGTC TTTGGCTCAG 4140
TTCAATTAAA AAGATATATCT ATTTGAAAGT TCTCAGAGT GTACATATGT TTCAGTATC 4200
AGGATCTGTA CATAAAAGTT TCTTCTCTAA ACCATTCAAC AAGAGCCAAT ATCTAGGCAT 4260
TTTCTTGTTA GCACAAATTT TCTTATTGCT TAGAAAAATG TCCTCCTTGT TATTCTGTT 4320
TGTAAGACTT AAGTACATTT GGTCTTTAAG GAAAGCAACG CTCCTCTGAA ATGCTGTGCT 4380
TTTTCTGTTT GCGGAAATAG CTGCTCTTTT TTCGGGAGTT AGATGTATAG AGTGTGTGTA 4440
TGTAACACTT TCTGTAGGAC ATCACCATGA ACAAAGATAT ATTTTCTATT TATTATTAT 4500
ATGTGCACTT CAAGAAGTCA CTGTGAGAGA AATAAAGAA TGTCTTAAAT GTCATGATTG 4560
GAGATGCTCT TTGCATTGCT TGGAAGGGGT GTACCTAGAG CCAAGGAAAT TGGCTCTGTT 4620
TTGAAAAAAT TTGCTGTTA TTATAGTAA CATACAAAGG ATGTCAAAA AAAAAAATAA 4680
AAAAAATAA AAAAAAATAA AA

```

30 Seq ID NO: 453 Protein sequence
Protein Accession #: Eos sequence

35
40
45

```

1 11 21 31 41 51
MGAAGRQDFL FKAMLTISWL TLTCPFGATS TVAAGCPDQS PELQPNPQGH DQDHHVHIQ 60
GKTLILLTSSA TVYSIHISGG GKLVIKDHDE FIVLRTRHIL IDNGGELHAG SALCPFGQNF 120
TILLYGRADE GIQPDPIYGL KYIGVGKGA LELHGQKLS WFLNKTLLHP GGMABGGYFF 180
ERSWGHGVI VHVIDPKSGT VIHSDRFDTY RSKKESERLV QYLAUVPDGR ILSVAVNDEG 240
SRNLDDMARK AMTKLGSXHF LHLGFRHPS FLTVKGNPSS SVEDHIEYHG HRGSAARVVF 300
KLFQTEHGYV FNVSLSEWV QDVETWEPD HDKVSQTKGG EKISDLWKAH PGKICNRPID 360
IQATTMDGVN LSTEVYKKG QDYRFACYDR GRACRSYVR FLCKGVPKPK LTVTIDTNVN 420
STILLNEDNV QSWKPGDILV IASTDYSMYQ AEEFOVLPGR SCAPNQKVA GKPMYLIHGE 480
EIDGVDMAE VGLLSRNIIV MGEMEDKCYR YRNHICNFPD PDTFGGHIKF ALGFKAAHLE 540
GTSLKHMGGQ LVQGYPIHFP LAGDVDERGG YDPPTYIRDL SIHHTFSRCV TVHSGNGLLI 600
KDVVGVNLSG HCFPTEDGPE ERNTFDHCLG LLVKSGLTLP SDRDSKMKCM ITEDSYPGYI 660
PKPRQCNNAV STFWMANPNN NLINCAAGS EETGFWFIH HVPTGPSVGM YSPGYSEHIP 720
LGKFYNNRAH SNRYGMIIID NGVKTTEASA KDKRPFLSII SARYSPHQA DPLKPREPAI 780
IRHFIAYKNQ DHGAWLRGGD VWLDSCHFRG EAQEGFLLTG MKAGGILLGG DEAAAGMAQG 840
FSPPCRCLLK LVTTGSPFAH VSLAHS

```

50 Seq ID NO: 454 DNA sequence
Nucleic Acid Accession #: NM_013282.2
Coding sequence: 85..2466

55
60
65
70
75
80
85

```

1 11 21 31 41 51
CGACTCCTTA GAGCATGGCA TGGCTCAGAG GTGCTGGTAA AACTGATGGG GGTTTTGTCT 60
GTCCCTCCCC TCAGCCCGCA CACCATGTGG ATCCAGGTTC GGACCATGGA CGGGAGGCAG 120
ACCCACAGCG TGGACTCGCT GTCCAGGCTG ACCAAGGTGG AGGAGCTGAG GCGGAAGATC 180
CAGGAGCTGT TCACAGTGGA GCCAGGCTG CAGAGGCTGT TCTACAGGGG CAACACAGATG 240
GAGGAGCGCC ATACCTCTCT CGACTACGAG GTCCGCCTGA ATGACACCAT CCAGCTCCTG 300
GTCCGCGAGA GCTCTGTGCT CCCCACAGC ACCAAGGAGC GGGACTCCGA GCTCTCCGAC 360
ACCGACTCCG GCTGCTGCTT GGGCCAGAGT GAGTCAGACA AGTCTCCAC CCACGCGCAG 420
GCGGCCGCGG AGACTGACAG CAGGCCAGCC GATGAGGACA TGTGGGATGA GACGGAATTG 480
GGGCTGTACA AGGTCAATGA GTACGTCGAT GCTCGGGACA CGAACATGGG GCGTGTGTTT 540
GAGGCGCAGG TGTCAAGGTT GACGCGGAAG GCCCCTCCCC GGGACGAGCC CTCGAGCTCC 600
ACGTCCAGCG CGCGCTGGA GAGGACGTC ATTTACCAAG TGAATAACGA CGACTACCCG 660
GAGAACCGCG TGTTCAGAT GAACTCCAG GAGTCCGAG CGCGCGCCG CACCATCATC 720
AAGTGGCAGG ACCTGAGGTT GGGCCAGGTG GTCATGCTCA ACTACAACCC CGACAACCCC 780
AAGGAGCGGG GCTTCTGGTA CGACGCGGAG ATCTCCAGGA AGCGCGAGAC CAGGACGGCG 840
CGGGAACACT ACGCCAACGT GGTGCTGGGG GATGATTCTC TGAACGACTG TCGGATCATC 900
TTGTTGACG AAGTCTTCAA GATTGAGCGG CCGGTGAAG GGAGCCCCAT GGTGACAA 960
CCCATGAGAC GGAAGAGCGG GCCGTCTGCG AAGCACTGCA AGGACGAGCT GAACAGACTC 1020
TGCCGGTCTT GCGCTGCTCA CTTGTGCGGG GCGCGGAGG ACCCGACAA GCAGCTCATG 1080
TGCGATGAGT GCGACATGGC CTTCCATATC TACTGCTTGG ACCCGCCCT CAGCAGTGT 1140
CCCAGCGAGG ACGATGGTA CTGCCCTGAG TGCCGGAATG ATGCCAGCGA GGTGTACTAG 1200
GCGGAGAGCG GCTGAGAGA GAGCAAGAAG AAGCGGAAGA TGCGCTCGCG CACATGCTCC 1260
TCACAGCGGG ACTGGGGCAA GGGCATGGCC TGTGTGGGCC GCACCAAGGA ATGTACATC 1320
GTCCCGTCCA ACCACTACGG ACCCATCCCG GGGATCCCGG TGGGCACCAT GTGGCGGTTC 1380
CGAGTCCAGG TCAGCGAGTG GGGTGTCCAT CGGCCCAAG TGCTGTGGAT ACACGGCGCG 1440
AGCAACGAGC GAGCGTACTC CTAAGTCTCT GCGCGGGGCT ATGAGGATGA CGTGACCAT 1500
GGGAATTTTT TCACATACAC GGTAGTGGT GGTGAGATC TTTCCGCAA CAAGAGGACC 1560
GCGGAACAGT CTGTGATCA GAAACTCACC AACACCAACA GGGCGCTGGC TCTCACTGC 1620
TTTGCTCCCA TCAATGACCA AGAAGGGGCC GAGGCCAAGG ACTGGCGTGC GGGGAAGCG 1680
GTGAGGGTGG TGGCAATGT CAAGGGTGGC AAGAATAGCA AGTACGCCCC CGCTGAGGGC 1740
AACCGCTACG ATGGCATCTA CAAGGTTGTG AATACTTGGC CCGAGAAGGG GAAGTCCGGG 1800
TTTCTGTTGT GCGCTTACTC TCTGCGGAGG GACGATGATG AGCCTGGCCC TTGACGGAAG 1860
GAGGGGAAGG ACCGGATCAA GAAGCTGGGG CTGACCATGC AGTATCCAGA AGGCTACCTG 1920

```

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

GAAGCCCTGG CCAACCGAGA GCGAGAGAAG GAGAACAGCA AGAGGGAGGA GGAGGAGCAG 1980
CAGGAGGGGG GCTTCGCGTC CCCCAGGACG GGCAGGGGCA AGTGAAGCGG GAAGTCGGCA 2040
GGAGGTGGCC CGAGCAGGGC CGGGTCCCGC CGCGGACAT CCAAGAAAAC CAAGGTGGAG 2100
CCCTACAGTC TCACGGCCCA GCAGAGCAGC CTCATCAGAG AGGACAAGAG CAACGCCAAG 2160
CTGTGGAATG AGGTCTCTGC GTCACTCAAG GACCGGCGGG CGAGCGGCAG CCGGTTCAG 2220
TGTTTCCTGA GTAAAGTGA GAGAGCGTTC CAGTGTATCT GCTGTCAGGA GCTGGTGTTC 2280
CGGCCCATCA CGACCGGTG CCAGCACAAC GTGTGCAAGG ACTGCCTGGA CAGATCCTTT 2340
CGGGCAGAGG TGTTTCAGTC CCCTGCCTGC CGCTACGACC TGGGCGCGAG CTATGCCATG 2400
CAGGTGAACC AGCCTCTGCA GACCGTCTCT AACCACTCT TCCCCGGCTA CGGCAATGGC 2460
CGGTATCTC CAAGCACTTC TCGACAGGCG TTTTGTCTGA AACGTGTGCG AGGGCTCGTT 2520
CATCGGCACAT GATTTTGTTC TTAGTGGGCT TAACCTAAAC AGGTAGTGT TTCTCGTTC 2580
CCTAAAAAGG TTTGTCTTCC TTTTITTTTA TTTTATTTT TCAATCTAT ACATTTTCAG 2640
GAATTTATGT ATCTGGCTA AAAGTTGGAC TTCTCAGTAT TGTGTTTAGT TCTTTGAAAA 2700
CATAAAAGCC TCGCAATTCT CGACAAAACA ACACAAGATT TTTTAAAGAT GGAATCAGAA 2760
ACTAGTGGT GTGGAGGCTG TTGATGTTTC TGGTGTCAAG TTCTCAGAAG TTGCTGCCAC 2820
CAACTCTTTA AGAAGGGCAG AGGATCAGTC CTCTCTAGG GTTCTGGCCC CCAAGGTGAG 2880
AGCAAGCATC TTCCTGACAG CATTTTGTCA TCTAAAGTCC AGTGACATGG TTCCCCGTGG 2940
TGGCCCGTGG CAGCCCGTGG CATGGCGTGG CTCAGCTGTC TGTGAAGTT GTTGCAAGGA 3000
AAAGAGGAAA CATCTCGGGC CTAGTTCAAA CCTTTGCCTC AAAGCCATCC CCCACAGAC 3060
TGCTTAGCGT CTGAGATCAAG CGTGAAAGT CCTCTGCCCA CGAGAGCAGG GAGTTGGGGC 3120
CAGCGAGAA TGGCTCAAG GGGACTCTGC TCCACGTGGG GCCAGGCGTG TGACTGACGC 3180
TGTCCGACGA AGGCGGCCAC GGACGGACGC CAGCACACGA AGTCACGTGC AAGTGCCTTT 3240
GATTCGTTC TTCTTCTAA AGACGACAGT CTGTGTGTT AGCACTGAAT TATTGAAAA 3300
GTCAACAGCA TACCTAGAAA TGCGGTCTAT CAGTCTCTCC TGACACCGGA TGGGTGCTG 3360
GGAACCGTT GAGCCTTATA GATCATTAC ATTCAATTT TTTAACTCAG CAAGTGAGAA 3420
CTTACAAGAG GGTTTTITTT TAATTTTTT TTCTCTTAAT GAACACATTT TCTAAATGAA 3480
TTTTTTTTGT AGTTACTGTA TATGTACCA GAAAGATATA ACGTTAGGCT TTGGTTGTTT 3540
TGTTTTTGT ATTTTITTT TTTTGAAAGG GTTGTGTAAT TTTTCTAATT TTACCAAGT 3600
TTGCAGGCTA TACCTCAATA AAACAGGGAT ATTTTAAATC ACATACCTGC AGACAAACTG 3660
GAGCAATGTT ATTTTAAAG GGTTTTITTC ACCTCCTTAT TCTTAGATTA TTAATGTATT 3720
ATGGGAAGAT GAGACAATT TGTGTAGGCT TTTTCTAAAG TCCAGTACTT TGTCCAGATT 3780
TTAGATTCTC AGAATAAATG TTTTTCACAG ATTGAAAAAA AAAAAAAA

Seq ID NO: 455 Protein sequence
Protein Accession #: NP_037414.2

1 11 21 31 41 51
MNIQVRTMDG RQTHTVDSL RLTKVEELRR KIQELFHVPE GLQRLFYRGK QMEDGHTLPD 60
YEVRINDTIQ LLVRSQSLVLP HSTKERDSEL SDTDSGCCLG QSESDKSSTH GEAAAETDSR 120
PADEDMWDET ELGLYKVNVE VDARDTNMGA WFEAQVVRVT RKAPSRDEPC SSTSRPALEE 180
DVIYHVKYDD YPENGUVQMN SRDVRARART IIKWQDLEVG QVVMNLNVPD NPKERGFWDY 240
AEISRRKRETR TARELYANVV LGDDSLNDCR IIFVDEVFKI ERPGEGSPMV DNPMMRKS GP 300
SCKHCKDDVN LKCRVACACH CGGRQDPDKQ LMCDECDMAF HIYCLDPPLS SVPSSEWYVC 360
PECNDADEV VLAGERLRES KKKAKMASAT SSSQRDNGKG MACVGRTEKC TIVPSNHYGP 420
IPGIFVGTMM RFRVQVSESG VHRPHVAGIH GRSDNGAYSL VLAGGYEDDV DHGNFFTYTG 480
SGGRDLGSKN RTABQSCDQK LNTNRLALAL NCFAPINDQE GAEAKDNRSG KPVVRVVRNVK 540
GGKNSKYAPA EGNRYDGIYK VVKYWPKEKG SGFLVWRYLL RRDDEDEPGW TKEGKDIRIK 600
LGLTMQYPPG YLEALANRER EKENSKEEES EQQEGGFASP RTGKGKWKRR SAGGGPSRAG 660
SPRRTSKKTK VEPYSLTAQO SSLIREDKSN AKLWNEVLAS LKDRPASGSP FQLFLSKVEE 720
TFQCICQCEL VFRPITTVQO HNVCKDCLDR SFRAQVFSFC ACRYDLGRSY AMQVNPQLQT

Seq ID NO: 456 DNA sequence
Nucleic Acid Accession #: NM_001200.1
Coding sequence: 325..1514

1 11 21 31 41 51
GGGGACTTCT TGAAGTTCGA GGGAGAATAA CTTGCGCACC CCACTTTGCG CCGGTGCGCTT 60
TGCCCCAGCG GAGCCTGCTT CGCCATCTCC GAGCCCCACC GCCCTCCAC TCCTCGGCCT 120
TGCCCCAGAC TGAGAGCGTG TTCCACGCGT GAAAGAGAG ACTGCGCGGC CGGCACCCGG 180
GAGAAGGAGG AGGCAAGAA AAGGAACGGA CATTCGGTCC TTGCGCCAGG TCCTTTGACC 240
AGAGTTTTC CATGTGGAGC CTCTTTCAAT GGACGTGTCC CGCGTGCCT CTTAGACGGA 300
CTGCGGTCTC TAAAGGTGCG ACCATGGTGG CCGGGAACCG CTGTCTTCTA GCGTTGCTGC 360
TTCCCCAGGT CCTCCTGGGC GGCGCGGCTG GCCTCGTTCG GAGCTGGGC CGCAGGAAGT 420
TCGCGGCGGC GTCGTGGGCG GCCTCCTCAT CCCAGCCCTC TGACGAGGTC CTGAGCGAGT 480
TCGAGTTGCG GCTGCTCAGC ATGTTGCGCC TGAACACAG ACCCACCCCC AGCAGGGAGC 540
CCGTGGTGCC CCCCTACATG CTAGACCTGT ATCGCAGGCA CTCAGGTGAG CCGGGCTCAC 600
CCGCCCCAGA CCACCGGTTG GAGAGGGCAG CCAGCCGAGC CAACACTGTG CGCAGCTTCC 660
ACCATGAAGA ATCTTTGGAA GAACACACAG AAACGAGTGG GAAACCAACC CGGAGATTCT 720
TCITTAATTT AAGTCTTATC CCACGGAGG AGTTTATCAC CTCAGCAGAG CTTCAGGTTT 780
TCCGAGAACA GATGCAAGAT GCTTTAGGAA ACAATAGCAG TTTCCATCAC CGAATTATA 840
TTTATGAAT CATAAAACCT GCAACAGCCA ACTCGAAAT CCCCGTGACC AGACTTTTGG 900
ACACCAAGTT GGTGAATCAG AATGCAAGCA GGTGGGAAAG TTTTATGTCT ACCCCGCTG 960
TGATGCGGTG GACTGCACAG GGACACGCCA ACCATGGATT CGTGGTGGAA GTGGCCCACT 1020
TGAGAGAGAA ACAGGTGTCT TCCAAGAGAC ATGTTAGGAT AAGCAGGTCT TTGCACCAAG 1080
ATGAACACAG CTGGTCACAG ATAAGGCCAT TGCTAGTAAC TTTTGGCCAT GATGGAAAAG 1140
GGCATCTCT CCACAAAGA GAAAAACGTC AAGCCAAACA CAAACAGCGG AAACGCCTTA 1200
AGTCAGCTG TAAGAGACAC CCTTTGTACG TGGACTTCAG TGACGTGGGG TGAATGACT 1260
GGATTGTGGC TCCCCGGGG TATCAGCCCT TTTACTGCCA CGGAGAATGC CCTTTTCTCT 1320
TGGCTGATCA TCTGAATCC ACTAATCATG CCATTGTTC GACGTGTGCT AACTCTGTTA 1380
ACTCTAGAT TCCTAAGGCA TGCTGTGTCC CGACAGAACT CAGTGCTATC TCGATGCTGT 1440
ACCTTGACGA GAATGAAAG GTTGATTATA AGAATATCA GGACATGGTT GTGGAGGGTT 1500
GTGGGTGTCG CTAGTACAGC AAAATTAAAT ACATAAATAT ATATATA

Seq ID NO: 457 Protein sequence
Protein Accession #: NP_001191.1

1 11 21 31 41 51
 5 MVAGTRCLLA LLLPQVLLGG AAGLVPELGR RKFAAASSGR PSSQPSDEVL SEFELRLLSM 60
 PGLRQRPTPS RDAVVPYML DLYRRHSQSP GSPAPDHRLE RAASRANTVR SFHHEESLEE 120
 LPETSGKTR RFFENLSSIP TEEFITSAL QVFRBQMQDA LGNNSSFHHR INIYEIKPA 180
 TANSKFPVTR LLD

10 Seq ID NO: 458 DNA sequence
 Nucleic Acid Accession #: NM_001999.2
 Coding sequence: 1..8736

1 11 21 31 41 51
 15 ATGGGGAGAA GACGGAGGCT GTGTCTCCAG CTCTACTTCC TGTGGCTGGG CTGTGTGGTG 60
 CTCTGGGCGC AGGGCAGCGC CGGCCAGCCT CAGCCTCCTC CGCCCAAGCC GCCCGGCCCC 120
 CAGCGCGCGC CGCAACAGGT TCGGTCCGCT ACAGCAGGCT CTGAAGCGCG GTTCTTAGCG 180
 CCCGAGTATC GCGAGGAGGG TGCCGAGGTG GCCAGCGCGC TCCGCGCGCG AGGACAGCAG 240
 GACGTGCTCC GAGGGCCCAA CGTGTGCGGC TCCAGATTCC ACTCCTACTG CTGCCCTGGA 300
 20 TGGAGAGCGC TCCCTGGAGG AAACCAAGTGC ATTGTCCCGA TTGTAGAAAT TAGTGTGGGA 360
 GATGGATTTT GTTCCCGTCC TAACATGTGT ACTTGTTCCT GTGGGCAAAAT ATCATCAACC 420
 TGTGGATCAA AATCAATTCA GCAGTGCAGT GTGAGATGCA TGAATGGTGG GACCTGTGCA 480
 GATGACCACT GCCAGTGCCA GAAAGGATAT ATTGGAACCT ATTGTGGACA ACCTGTCTGT 540
 GAAAATGGAT GTCAAGATGG TGGACGTTCG ATCGCCCAAC CGTGTGCTGT TGTATTATGGG 600
 25 TTCACTGGTC CACAGTGTGA AAGAGATTAC AGGACAGGCC CGTGTTCAC TCAGGTCAAC 660
 AACCAATGTG GCCAAGGGCA GCTGACAGGC ATTGTCTGCA CGAAGACTCT GTGCTGTGCC 720
 ACCACTGGAC GGGCGTGGGG CCATCCCTGT GAGATGTGTC CAGCCAGGCC TCAGCCCTGC 780
 CGACGGGGTT TCATCCCAA CATCCGCACT GGAGCTTGCC AAGATGTGTA TGAATGCCAG 840
 GCTATCCAGG GGTATGCCA AGGAGGAAAC TGTATCAATA CAGTGGGCTG TTTTGAATGC 900
 30 AGATGCCCTG CTGCTCACA ACAGAGTGAA ACTACTCAGA AATGTGAAGA CATGTATGAG 960
 TGCAGATCA TTCTCTGGAT ATGTGAACT GGTGAATGTT CCAACACCGT GGGAAAGCTAT 1020
 TTTTGTGTTT GTCCAGTGG ATATGTAACC TCAACAGATG GCTCTGATG CATCGATCAG 1080
 AGAACAGGCA TGTGTTTCTC GGGCCTGGTG AATGGCCGCT GTGCACAAGA GCTCCCGGGG 1140
 AGAATGACGA AATGTCAAGT CTGCTGTGAG CTTGGCCGCT GCTGGGCGAT CGGAACCAT 1200
 35 CCTGAGCCTG GTCTGTGAG AGGTTCTGAG GAAATCGCA GACTTGCAT GGATGGACTT 1260
 CCAATGGGAG GAATTCGAG GAGTGTGGT TCCAGACCTG GAGGCACTGG GGGAAATGGC 1320
 TTTGCCCAA GTGGCAATGG CAATGGCTAT GGGCCAGSAG GGACAGGCTT CATCCCATC 1380
 CCTGGAGGCA ATGGCTTTTC TCCTGGCGTT GGGGAGCGG GTGTGGGGGC CGGGGGACAG 1440
 GGACCTATCA TCATGGAAT AACAATTCTG AACAGACAA TAGATATCTG TAAGCATCAT 1500
 40 GCTAACCTTT GTTAAATGG ACCTGTGATA CCAACTGCTC CAGACTACCG ATGTGAATGC 1560
 AACATGGGTG ATAAGCAGGA TGCAATGGA GATTGTATAG ATGTGTATGA ATGCACATCA 1620
 AATCCCTGCA CTAATGGAGA TTGTGTTAAC ACACCTGGTT CCTATTATG TAAATGTCAT 1680
 GCTGGATTCC AGGAGGACTCC TACCAAGCAA GCATGCATTG ATATTGATGA GTGCATCCAG 1740
 45 AATGGGGTTC TTGTAAAAA CGGTGATGTC GTGAACCTG ATGGAAGTTT CCAGTGCATT 1800
 TGCAATGCGG GCTTGAATTT AACTACAGAT GGAAGAAACT GTGTGATCA TGATGAATGT 1860
 ACAAATGAGC GAATGTGTTT GAATGGAATG TGCTCAATG AAGATGGCAG CTTCAAGTGC 1920
 ATCTGCAAAC CAGGATTTGT CTTGGCTCCA AATGGGCGTT ACTGTACTGA TGTGTATGAA 1980
 TGCCAGACCC CAGGAATCTG CATGAATGGG CACTGCATCA ACAGTGAAGG GTCTTCCCG 2040
 50 TGTGACTGTC CCCAGGCGCT GGCTGTGGGC ATGATGATGG GTGTGTGTGT TGATACTCAC 2100
 ATGCGCAGTA CCTGCTATGG AGGAATCAAG AAAGAGAGTGT GTGTGCGTCC TTTCCCGGTT 2160
 GCAATGACCA AGTCCGAATG CTGCTGTGCC AATCCAGACT ATGCTTTTGG AGAACCTTGC 2220
 CAGCCATGCC CTGCAAAAAA TTCAGCTGAA TTCCACGSCC TTTGTAGTAG TGGAGTAGGT 2280
 ATCACTGTGG ATGGAAGAGA TATCAATGAA TGTGCTTTGG ATCTGATAT ATGTGCCAAT 2340
 55 GGGATTGTG AAAACTTACG TGGTAGTTAC CGTTGTAATT GCAACAGTGG CTATGAACCA 2400
 GATGCTCTG GAAGAACTG TATTGACATT GATGAATGTT TAGTAAACAG ACTGCTTTGT 2460
 GATAAGCGAT TGTGCGGAAA CAGCCAGGA AGTTACAGCT GTACGTGCC ACCAGGTAT 2520
 GTGTTCAAGA CTGAGACAGA GACCTGTGAA GATATAAATG AATGTGAAG CAACCCATGT 2580
 GTCAATGGGG CCTGAGAAA CAACCTTGGA TCTTTCAATT GTGAATGTTT GCCCGGACG 2640
 60 AAATCAGCT CCAAGGATT GATCTGTATT GACAGCCTGA AGGGGACCTG TTGGCTCAAC 2700
 ATCCAGGACA CGCGCTGTGA GGTGAATATT AATGGAGCCA CTCTGAAATG TGAATGCTGT 2760
 GCCACCTCG GAGCGCGCTG GGGGAGCCCC TGTGAGCGGT GTGAACCTGA TACAGCTTGC 2820
 CCAAGAGGGG TTGCCAGGAT TAAAGGTGTT ACCTGTGAAG ATGTTAATGA GTGTGAGGTG 2880
 TTTCTTGGG TTGTCCAAA TGGACGCTGT GTCAACAGTA AGGGATCTTT TCATTGCGAG 2940
 65 TGCCCTGAAG GCCTTACGTT GGATGGGACT GGCGGTGTAT GTTTGGATAT TCGCATGGAG 3000
 CAGTGTACT TGAAGTGGGA TGAAGATGAA TGCAATCCAC CCGTTCCTGG AAAGTTCCGC 3060
 ATGGATGCCCT GTGCTGTGTC TGTGCGGGCG GCTTGGGGCA CCGAGTGTGA GGAGTGCCCC 3120
 AAACCTGGCA CCAAGGAATA CGAGACACTG TGCCCCCGCG GGGCTGGCTT TGCTAACCGA 3180
 GGGGATGTTT TTAAGTGGCG GCCATTTTAC AAAGACATCA ATGAATGCAA AGCAATTTCT 3240
 70 GGGATGTGCA CTTATGGGAA GTGCAGAAAT ACAATCGGAA GCTTCAAATG CCGTTGCAAT 3300
 AGTGGCTTTG CTCTAGACAT GGAGGAAAGA AACTGCACGG ACATCGACGA GTGCAGGATT 3360
 TCTCTGACCT TGTGTGGCAG TGGAACTGCG GTCAATACAC CGGGCAGCTT TGAGTGGCAG 3420
 TGCTTGAAG GCTATGAAG TGGCTTCATG ATGATGAAGA ACTGCATGGA CATGACGGA 3480
 TGTGAACGTA ACCCTCTCCT TTGTAGGGGT GGCACCTGTG TGAACACTGA GGGCAGCTTT 3540
 75 CAGTGTGACT GCCCACTGGG ACACGAGCTG TCACCATCCC GTGAGGACTG TGTGGATATT 3600
 AATGAATGCT CCTCAGTGA CAATCTCTCG AGAAATGGAA AATGTGTGAA CATGATTGGA 3660
 ACCTATCAGT GCTCTGTCAA TCCTGGATAT CAGGCTACGC CAGACCGCCA GGGCTGTACA 3720
 GATATTGATG AATGTATGAT AATGAACGGA GGCTGTGACA CCCAGTGCAC AAATTCAGAG 3780
 GGAAGCTACG AATGCAGCTG CAGTGAAGGT TATGCCCTGA TGCCAGATGG GAGATCGTGT 3840
 80 GCAGACATTG ATGAATGTGA AAACAATCCT GATATCTGTG ATGCGCGCCA GTGTACCAAC 3900
 ATTCTGAGAG AGTATCGCTG CCTCTGCTAT GATGGCTTCA TGGCTTCCAT GGACATGAAA 3960
 ACATGCAITG ATGTCAATGA ATGTACCTA AATTCAAATA TCTGCATGTT TGGGGAATGT 4020
 GAGAACACAA AGGGATCCTT CATTTGCCAC TGTACAGTGG GTTACTCAGT GAAGAAGGGG 4080
 ACCACAGATG GTACAGATGT GGATGAGTGT GAAATGGGTG CTCATAACTG CGACATGCAT 4140
 85 GCCTCATGTC TGAATATCCC AGGAAGCTTC AAGTGTAGCT GCAGAGAAGG CTGGATTGGA 4200
 AACCGCATCA AGTGTATTGA TCTGGACGAA TGTCTAATG GAACCCACCA GTGTAGCATC 4260
 AATGCTCAGT GTGTAATATC CCGGGCTCA TACCGCTGTG CCGTCTCGGA AGGTTTCACT 4320
 GGTGATGGCT TTACCTGCTC AGATGTTGAT GAGTGTGCGA AAAACATAAA CCTCTGTGAG 4380

	AACGGACAGT	GCCTTAATGT	CCCGGGTGCA	TATCGCTGCG	AGTGTGAGAT	GGGCTTCACT	4440
	CCAGCCTCAG	ACAGCAGATC	CTGCCAAGAT	ATTGATGAAT	GCTCCTTCCA	AAACATTGTG	4500
	GTCTCTGGAA	CATGTAATAA	CCTGCCTGGA	ATGTTTCATT	GCATCTCGGA	TGATGGTTAT	4560
5	GAATTGACAA	GAACAGGAGG	GAACGTGTACA	GATATTGATG	AGTGTGAGAT	TCCTATAAAC	4620
	TGTTGCAATG	GCCTATGTGT	CAACACGCCT	GGTCCGTATG	AGTGTAACTG	CCCACCCGAT	4680
	TTTCAGTTGA	ACCCAACTGG	TGTGGGTTGT	GTGACAAACC	GTGTTGGGCA	CTGCTACCTG	4740
	AAGTTTGGAC	CTGAGGAGGA	TGGGAGTCTG	TCTTGCAACA	CCGAGATCCG	GGTGGGCGTC	4800
	AGTGGCTCTT	CATGCTGCTG	CTCTCTGGGA	AAGGCCTGGG	GAACCCCTCG	TGAGACATGC	4860
	CCCGCTGTGA	ATAGCACTGA	ATATTACACC	CTGTGTCCCG	GAGGTGAAGG	CTTCAGACCT	4920
10	AACCCCATCA	CAATCATTTT	AGAAGACATT	GACGAATGCC	AGGAGTTACC	AGGTCTCTGC	4980
	CAGGTTGGAA	ACTGCATCAA	CACCTTTGGG	AGCTTCCAGT	GTGAGTGCCC	ACAAGGCTAC	5040
	TACCTCAGCG	AGGATACCCG	CATCTGTGAG	GATATTGATG	AGTGTTTTGC	ACATCCTGGT	5100
	GTGTGTGGGC	CTGGGACCTG	CTATAACACC	CTGGGAAATT	ACACCTGCAT	TGCCCCACCT	5160
	GAGTACATCG	AGGTCAATGG	AGGCCACAAC	TGCATGGACA	TGAGAAAAAG	CTTTTGTCTC	5220
15	CGAAGCTATA	ATGGAACCCAC	TTGTGAGAAT	GAGTTGCCCT	TCAATGTGAC	AAAAAGGATG	5280
	TGCTGTGCGA	CATATAATGT	GGGCAAGACT	GGGAACAAAC	CTTGTGAACC	ATGCCCAACT	5340
	CCAGGAACAG	CTGACTTTAA	AACCATATGT	GGAAATATTC	CTGGATTACG	CTTTGACATT	5400
	CACACAGGAA	AAGCTGTGTA	CATTGATGAA	TGTAAAGAGA	TTCCAGGCAT	TGTGCAAAAT	5460
	GGTGTGTGCA	TAAACAGAT	TGGCAGTTTC	CGCTGTGAAT	GCCCTACAGG	ATTCACTTAC	5520
20	AATGACCTGC	TGTTGGTTTG	TGAAGATATA	GATGAGTGCA	GCAATGGTGA	TAATCTCTGC	5580
	CAGCGGAATG	CAGACTGCAT	CAATAGTCCT	GGTAGTTACC	GCTGTGAATG	TGCCGCGGGT	5640
	TTCAAACTTT	CACCAATGGT	GGCCTGTGTA	GATCGCAATG	AATGTTTAGA	AATTCTTAAC	5700
	GTGTGCAATG	ATGGCTTTGT	TGTTGATCTG	CAAGGAAGTT	ACCACTGCAT	CTGCCCAAT	5760
	GGCTTTAAGG	CTTCTCAGGA	CCAGACCATG	TGCATGGATG	TTGATGAGTG	CGAGCGGCAC	5820
25	CCATGTGGAA	ATGGAACCTG	TAAAAACACC	GTGGATCCCT	ATAACTGTCT	GTGCTACCCA	5880
	GGGTTTGAAC	TCACATATAA	TAATGATTGC	CTGGACATAG	ATGAGTGACG	TTCTTTTFTT	5940
	GGTCAGGTGT	GCAGAAATGG	ACGTTGTTTT	AATGAAATGG	GTTCTTTCAA	GTGCTATATG	6000
	AACGAAGGTT	ATGAACCTAC	CCCAGATGGC	AAAACTGTGA	TAGACACTAA	TGAGTGTGTC	6060
	GGCCTTCCCG	GCTCTTGCTC	TCTTGGTACC	TGTCAGAAAT	TGGAGGGATC	CTTCAGATGC	6120
30	ATCTGTCCCG	CAGGGATAGA	AGTAAAAAGC	GAGAACTGCA	TTGATATAAA	TGAATGTGAT	6180
	GAAGATCCCA	ACATTTGTCT	TTTTGGTTCC	TGTACTAATA	CTCCAGGGGG	CTTCCAGTGC	6240
	CTCTGCCCTG	CTGGCTTTGT	ACTATCTGAT	AATGGACGGA	GATGCTTTGA	TACTCGCCAG	6300
	AGCTTCTGCT	TCACAAATTT	TGAAAATGGA	AAGTGTCTG	TACCCAAAGC	TTTCAACACC	6360
	ACAAAAGCAA	AATGCTGCTG	TAGTAAGATG	CCAGGAGAGG	GCTGGGGGGA	CCCCTGTGAG	6420
35	CTGTGCCCCA	AAGACGATGA	AGTTGCATTT	CAGGATTGTT	GTCCATATGG	CCATGGAACT	6480
	GTCCCTAGTC	TTATGATATC	ACGTGAAGAT	GTCATGAGT	GTCTTGAGAG	CCCAGGCATT	6540
	TGTTCAAATG	GTCATGTATG	CAACACCGAC	GGATCTTTTC	GCTGTGAATG	TCCAATGGGC	6600
	TACAACCTTG	ACTACACTGG	AGTACGCTGT	GTGGATACTG	ATGAGTGTTC	AATCGGCAAT	6660
	CCGTGTGGAA	ATGGTACATG	CACCAATGTT	ATTGGGAGTT	TTGAATGCAA	TTGCAATGAA	6720
40	GGCTTTGAGC	CAGGGCCCAT	GATGAATTTG	GAAGATATCA	ACGAATGTGC	CCAGAACCCA	6780
	CTGCTGTGTG	CTTTACGCTG	CATGAACACT	TTTGGTCCCT	ATGAATGCAC	GTGCCCGATT	6840
	GGCTATGCCG	TCAGGGAAGA	TCAAAAGATG	TGCAAAAGATC	TGGATGAATG	TGCTGAAGGG	6900
	TTACAGCACT	GTAATCTTAG	GGGCATGATG	TGTAAGAATC	TAATCGGCAC	CTTCATGTGC	6960
	ATCTGCCCTC	CTGGAATGGC	CGAAAGGCC	GATGGAGAAG	GCTGTGTAGA	TGAAAATGAA	7020
45	TGCAGGACCA	AGCCAGGAAT	CTGTGAAAT	GGACGTTGTT	TTAACATTAT	TGGAAGCTAT	7080
	AGATGTGAGT	GTAATGAAGG	ATTCCAGTCA	AGTTCTTCAG	GCACCTGAATG	CCTTGACAAT	7140
	CGACAGGGTG	TCTGCTTTGC	AGAGGTAAGT	CAGACAAAT	GTCAAATGGC	ATCCAGTAGT	7200
	CGCAATCTCG	TCACTAAGTC	AGAATGCTGC	TGTGATGGTG	GGCGAGGCTG	GGGCCACCAAG	7260
	TGGGAGCTTT	GGCCACTTCC	TGGAACCTGC	CAGTACAAAA	AGATATGTCC	TCAATGGCCA	7320
50	GGATATACAA	CTGATGGAAG	AGATATTGAT	GAATGTAAAG	TAATGCCAAA	CCTCTGCACC	7380
	AATGGTCAGT	GCATCAATAT	CATGGGCTCA	TTCCGATGCT	TCTGCAAGGT	TGGCTACACC	7440
	ACAGACATCA	TGGAACCTTC	TTGTATAGAC	CTTGATGAAT	GCTCCAGTGC	CCCGAAACCA	7500
	TGCAACTACA	TCTGCAAGAA	CACCTGAGGG	AGTTATCAGT	GTTCAATGCC	GAGGGGGTAT	7560
	GTCTGTCAAG	AGGATGGAAT	GACATGCAAA	GACCTTGATG	AATGTCAAAC	AAAGCAGCAT	7620
55	AACGTGCCAGT	TCCTCTGTGT	CAACACCTCG	GGGGGGTTTA	CCTGTAAATG	TCCACCTGGT	7680
	TTCAACAGCG	ATCAACACTG	TTGTATCGAC	AACAACGAAT	GTGGGTCTCA	ACCTTTGCTT	7740
	TGTGGAGGAA	AGGGAATCTG	TCAAAACACT	CCAGGCAGTT	TCAGCTGTGA	ATGCCAAAGA	7800
	GGGTCTCTCT	TTGATGCCAC	CGGACTGAAC	TGTGAAGATG	TTGATGAATG	TGATGGGAAC	7860
	CACAGGTGCC	AACACGGCTG	CCAGAACATC	CTGGGTGGCT	ACAGATGTGG	CTGCCCCCAA	7920
60	GGCTACATCC	AGCACTACCA	GTGGAATCAG	TGTGTGATG	AGAATGAATG	CTCCATATCC	7980
	AATGCCCTGT	GCTCTGCTTC	CTGCTACAAC	ACCTTGGGGA	GTTACAAGTG	CGCCTGCCCG	8040
	TCCGGGTTCT	CTCTGACCCA	GTTCTCCAGT	GCCTGCCACG	ACGTGAATGA	GTGCTGCTCC	8100
	TCCAGAAACC	CCTGCAATTA	CGGCTGCTCT	AACACGGAGG	GGGGCTACCT	CTGTGGCTGC	8160
	CCCCCTGGGT	ATTACAGAGT	GGGACAAGGC	CACGTGTGCT	CAGGAATGGG	ATTTAACAAG	8220
65	GGGCAGTACC	TGTCATCGGA	TACAGAGGTC	GATGAGGAAA	ATGCTCTGTC	CCCAGAAGCA	8280
	TGCTACAGAT	GCAAAATCAA	CGGCTATCCT	AAGAAAGACA	GCAGGCAGAA	GAGAAGTATT	8340
	CATGAACCTG	ATCCCACTGC	TGTTGAACAG	ATCAGCCTAG	AGAGTGTGCA	CATGGACAGC	8400
	CCCGTCAACA	TGAAGTTCAA	CCTCTCCAC	CTCGGCTCTA	AGGAGCACAT	CCTGGAACCTA	8460
	AGGCCCGCCA	TCCAGCCCTT	CAACAACCAC	ATCCGTTATG	TCATCTCTCA	AGGGAACGAT	8520
70	GACAGGCTCT	TCGGCATCCA	CAAAAGGAAT	GGGCTCAGCT	ACTTGACAC	GGCCAAGAAG	8580
	AAGCTCATGC	CCGGCACATA	CACACTGGAA	ATCACTAGCA	TCCTCTCTCA	CAAGAAGAAG	8640
	GAGCTTAAGA	AACCTGAAGA	GAGCAATGAG	GATGACTACC	TCCTAGGGGA	GCTTGGGGAG	8700
	GCTCTCAGAA	TAGGCTGCA	GATTGAGTCT	TATTAACCGT	TCACAGACTT	GGGCCACAGC	8760
	TCAAACTCTA	GCACAGCCAG	CTCTGAGAAG	CATTGAAAA	GTCAGGACT	AATTTTAAAG	8820
75	AGGAAAAATA	ATAATAACTC	TGTTTCTTTT	CCTCCCTGTC	TTAGACTTTG	AATGTTGACC	8880
	CTCACAGGGA	GGGATAATTT	AGACTCTGTG	ATGGCCAAAG	ATTTGAGCTC	AAAGGCCAAC	8940
	GTGGTTACTG	TATTTTCTAT	ATAACTTCAT	TTTAAATAT	ATTAAAGAA	ACCTAAATGT	9000
	TCAAGATATC	AGCATATGGC	ACTAAATGCA	CAAAATAAT	GTGAGCTTTT	TTTTTTTTTT	9060
	CCTGTTAGCA	GTCTGTAACA	CTTTGGGTAT	TTTGCTATAG	TTGCTAATTA	AAAAATATA	9120
80	GATGTTTATT	TATTTTAAAT	GCAGTAATAT	ATGGAGAAAT	GAACAACTA	TGTAACAAA	9180
	AAGGGAACCT	CACCTGTTTT	TCTTTAGATT	TATAAATTTG	AGCTATTTTT	TTTAGAGGTG	9240
	CTTTTAAATA	ATCCAAATAG	TACAAGAGAT	GTTTCTTTTG	GTTTTCTGCC	AGTCATCCAG	9300
	CTGATACACA	CCTGATCGAT	TTTAAAGAAA	GCCACACAGA	GCTGAATCGG	CGAGTGCTAA	9360
	TCAATAATTT	AAAAGACATG	AATGTCATTA	GATCCTTTAT	AACGTAGATC	GAAGCCAAAG	9420
85	CAGCTCATTT	GTGACAACT	TTTATATCAC	CAGACACACC	AGGCAACAGA	AGTTGAAGCA	9480
	CAACCACTGT	AGCAAAATAC	CTTGACTGCT	TGTGAGACCA	TTAGCACTGC	AGGCCAAACC	9540
	GTACTGTATT	TCCTTCTCAT	AACCTCAAGG	AACCATATGT	GCTACCCACA	ACACCTCATT	9600

WO 02/086443

PCT/US02/12476

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

```

CITACCCAGG GTGCGCTGCG TCCTCATGGT ACTGTAGGCA GCTGAAGAAC GCCTGTTCCC 9660
TTGAAAGGGA ACACCTGGCA TTCTGTGGTG TTCTGCTGCTG TCTTAAATAA TGGTGCATTT 9720
ATTATGTTCA AGTTATTTC A GATTGCCAT ATGTGCAAAAC AAATCATGCA ATGCAGCCAA 9780
GGAATATATG TTGTGTGTTG TTGTTTAAAC CCATTTTTTT TTTAGAATTT TCATTAATAC 9840
TGTAGTTATA CACCATATGC CTCATTTTAT CATAGCCTAT TGTGTATGAA AGATGTTTGT 9900
ACAATGAATT GATGTTTAGT TTGCTTTAGT CATTTAAAAA GATATTGTAC CAGGATGTGC 9960
TATTAAGAGC ACGTATCCAT TATTCTTCTC AACCCAAGAA CCTGTTTCTT GGACCACTGA 10020
CCAAACCTCA TATGTGAAAT GGCCAAAGCA CATGCAGGCT CCTGGTGTGT CCTCTCAAA 10080
CTGTGCTGAC CAAAGATTAG TAACCAAGTA TACCCAGTAT TTTGAGGTTT TATTGTTTT 10140
TTAATAACTA AAAAAAACT CGTGCC
  
```

Seq ID NO: 459 Protein sequence
Protein Accession #: NP_001990.1

15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

```

1 11 21 31 41 51
MGRRRRLCQ LYFLWLGCVV LWAQGTAGQP QPPPPKPPRP QPPPPQVRS A TAGSEGGFLA 60
PEYREEGAAY ASRVRRRGQQ DVLRGPNVCG SRFHYSQCPG WKTLPGGNQC IVPICRNSCG 120
DGFCSRFNMC TCSGGTSSST CGSKSIQCCS VRCMGGTCA DDHCQCKGY IGTYCGQFVC 180
ENGCGNGGRC IAQPCACVYG FTGPQCDERD RTGCPCTQVN NMCQCGQLTG IVCTKTLCCA 240
TTGRAGWGPC EMCPAQPPQC RRGFPINIRT GACQDVBQO AIPGICQGGN CINTVGSFEC 300
RCPAGHKQSE TTQKCEDIDE CSIIIPGICET GECSNTVGSY FCVCPRGVVT STDGSRCIDQ 360
RTGMCFSGLV NGRCAQLPQG RMTKMCCCE PGRCWGIGTI PEACPVRSSE EYRRLCMDGL 420
PMGGIPGSAG SRPGGTGGNG FAPSGNGNGY GPGGTGFIPI PGGNGFSPPV GGAGVGAGGQ 480
GPIITGLTIL NQTDIDCKHH ANLCLNGRCI PTVSSYRCEB NMGYKQDANG DCIDVDECTS 540
NPCTNGDCVN TPGSYCKCK AGFQRTPTKQ ACIDIDEICQ NGVLCKNGRC VNSDGSFQCI 600
CNAGFLTTTD GKNVCVHDCE TTNMCLNGM CINEDGSFKC ICKPFPVLAP NGRYCTDVDE 660
CQTPGICMNG HCINSEGSFR CDCPGLAVG MDGRVCDVTH MRSTCYGGIK KGVCVRPFP 720
AVTSECCCA PDYVGPEGPC QPCPAKNSAE FHGLCSSGVG ITVDGRDINE CALDPIDCAN 780
GICENLRGSY RCNCSNGYEP DASGRNCIDI DECLVNRLLC DNGLCRNTPG SVSCTCPPGY 840
VFRTEETCE DINECESNPC VNGACRNNG SFNCECSPGS KLSSTGLICI DSLKGTCHLN 900
IQDSRCEVNI NGATLKSECC ATLGAAWGS P CERCELDTAC PRGLARIKV TCEVNECEV 960
FFGVCPNGRC VNSKGFHCE CPEGLTLDGT GRVCLDIRME QCYLKWDEDE CIHPVPKFR 1020
MDACCAVGA AWGTECEEC KPGTKEYETL CPRGAGFANR GDVLTGRFFY KDINECKAF 1080
GMCTYKCRN TIGSFKRCRN SGFALDMEER NCTDIDECRI SPDLCSSGIC VNTPGSFEC 1140
CFEGYSEGF MKNKMDIDG CERNPLLCRG GTCVNTGESP QDCDPLGHEL SPSREDCVDI 1200
NECSLSDNLC RENGCVNMIG TYQCSNPGY QATPDRQCT DIDECEMIMG GCDTQCTNSE 1260
GSYECSSCG YALMPDGRSC ADIDECENNP DICDGGQCTN IFGEYRCLCY DGFMASMDMX 1320
TCIDVNECLD NSNIMRGEC ENTKGSFICH CQLGYSVKKG TTGCTDVDEC EIGAHCNDMH 1380
ASCINIPGSP KCSREGWIG NGIKCIDLDE CSNGTHQCSI NAQCVNTPGS YRCACSEGFT 1440
GDGFTCSVD ECAENINLCE NGQCLNVPGA YRCECEMGFT PASDSRSCQD IDECSFONIC 1500
VSGTCNLPD MFHICDDGY ELDRTGNCIT DIDECAPIIN CVNGLCVNTP GRVYCNCPD 1560
FQLNPTGVCG VDNRVGNCYL KFGPRGDSGL SCNTEIGVGV SRSSCCSLG KAWGNPCETC 1620
PFVNSTEYIT LCPGEGGFRP NPTIILEDI DECQLPGLC QGGNCINTFG SFQCECPQGY 1680
YLSEDTRICE DIDECAHPG VCGPOTCYNT LGNYTCICPP EYMQVNGGHN CDMRKSFCY 1740
RSYNGTTCEN ELFPNVTKRM CCCTYNVGKA GNKPCPCPT PGTADEFKTY GNIPGFTFDI 1800
HTGKAVDIDE CKNCVIGICAN GVCINQIGSF RCECPTGPSY NDLLLVCEDI DECSNGDNL 1860
QRNADCSNP GSYRCECAAG PKLSPNGACV DRNECLEIPN VCSHGLCVDI QGSYQCTICH 1920
GFKASQDTM CWDVECEERH PCNGTCKNT VGSYNLCYR GFELTHNND LDIDECSFF 1980
GQVCRNGRC NEIGSFKCLC NEGYELTPDG KNCIDTNECV ALPGSCSPGT QNLEGSFRC 2040
ICPPGYEVKS ENCIDINECD EDPNICLFGS CTNTPGGPQC LCPPGFVLSD NGRRCFDTRO 2100
SFCPTNFENG KCSVPKAPNT TKAKCCSKM PEGGWGDPCE LCPKDEVAEP QDLCPYHGT 2160
VPSLHDTRED VNECLESPI C SNGQCINTD GSFRCECPMG YNLDTGVRC VDTDECSIGN 2220
PCNGTCTNV IGSFECNCE GEPEGPMNC EDINECAONP LLALRCMNT FGSYEECTPI 2280
GYALREDQRM CKDLDECAEG LHDCESRGM CKNLIGTFMC ICPGMAARRP DGECCVDENE 2340
CRTKPGICEN GRVCNIGSY RCECEGFPQS SSSGTECLDN RQGLCFEVL QTIQMASS 2400
RNLVTSECC CDGGRGWGHQ CELCPLPGTA QYKKICPHGP GYTIDGRDID BCKVMPNLCT 2460
NGQCINTMGS FRFCFKVGT TDISGTSID LDECSQSPKP CNYICKNTEG SYQCSCPRGY 2520
VLQEDGKCTK LDDEQCTKH NCQFLCVNTL GGFTCKCPPG FTQHTTACID NNECGSPL 2580
CGGKICQNT PGFSCECQR GFSLDATGLN CEDVDECDGN HRCQHGQNI LGGYRCGCPQ 2640
GYIQHYQWNO CVDENCSNP NACGSASCYN TLGSYKCAP SGFSFDQFS ACHDVNECSS 2700
SKNFCYNGS NTEGGLCG PPGYRVGQG HCVSGMGFNK GQYLSLDETE DEENALSPEA 2760
CYECKINGYP KKSQRKRSI HEPDPTAVEQ ISLESVDMS PVMKFNLSH LGSKEHILEL 2820
RPAIQPLNKH TRVISQND DSVFRIHQRN GLSVLHTAK KIMPPTYTLE ITSIPLYKK 2880
ELKKLESNE DDYLLGELGE ALMRRLQIQ Y
  
```

Seq ID NO: 460 DNA sequence
Nucleic Acid Accession #: NM_013372.1
Coding sequence: 63..617

70
 75
 80
 85

```

1 11 21 31 41 51
GCGGCCGAC TCAGCGCCAC GCGTCGAAAG CGCAGGCCCC GAGGACCGGC GGCAGTACA 60
GTATGAGCGG CACAGCCTAC ACGGTGGGAG CCTGCTTCT CCTCTGGGG ACCCTGCTGC 120
CGGCTGCTGA AGGAAAAAG AAAGGTCCC AAGGTGCCAT CCCCCGCCA GACAGAGCCC 180
AGCACATGA CTCAGAGCAG ACTCAGTGC CCGAGCAGC TGGCTCCAG AACCAGGGG 240
GGGCAAGG GCGGGCACT GCCATGCCG GGGAGGAGT GCTGGAGTC AGCCAAGAG 300
CCCTGCATGT GCGAGAGCG AAATACCTGA AGCGAGACT GTGCAAAAC CAGCCGCTTA 360
AGCAGACCAT CACAGGAA GCCTGCAACA GTCGCACCAT CATCAACCG TCTGTATAG 420
GCCAGTGCAA CTCTTTCTAC ATCCCCAGG ACATCCGAA GGAGGAAGT TCCTTTTCA 480
CCTGCTCCTT CTGCAAGCCC AAGAAATTCA CTACCATGAT GGTCACTC AACTGCCTG 540
AACTACAGCC ACCTACCAAG AAGAAGAGAG TCACACGTGT GAAGCAGTGT CGTTGCATAT 600
CCATCGATT GATTAAAGCC AATCCAGGT GCACCCAGCA TGCTCTAGGA ATGCAGCCCC 660
AGGAAGTCCC AGACCTAAA CAACCAGAT CTTACTTGGC TTAACCTAG AGGCCAGAAG 720
AACCCGAGC TGCTCTCTG CAGGAGCCTG CTTGTGCGTA GTTCGTGTG ATGAGTGTG 780
ATGGGTGCT GTGGGTGTT TTAGACACCA GAGAAACAC AGTCTCTGT AGAGAGCACT 840
CCCTATTGTT TAAACATATC TGCTTTAATG GGGATGTACC AGAAACCCAC CTCACCCCG 900
  
```


CTCACATCTA AAGGGGCGGG GCGTGGTCT GGTCTGACT TTGTGTTTT GTGCCCTCCT 960
 GGGGACCAGA ATCTCCTTTC GGAATGAATG TTCATGGAAG AGGCTCCTCT GAGGGCAAGA 1020
 GACCTGTTTT AGTGTCTGAT TCGACATGGA AAAGTCCTTT TAACTGTGTC TTGCATCTCT 1080
 CTTTCTCTCT CTTCTCTACA ATCCATCTCT TCTTAAGTTG ATAGTGACTA TGTCAGTCTA 1140
 ATCTCTGTGT TGCCAAAGTT CCTAAATTAA TTCACITTAAC CATGATGCAA ATGTTTTTCA 1200
 TTTTGTGAAG ACCCTCCAGA CTCTGGGAGA GGCTGGTGTG GGCAAGGACA AGCAGGATAG 1260
 TGGAGTGAGA AAGGGAGGGT GGAGGGTGAG GCCAAATCAG GTCCAGCAAA AGTCAGTAGG 1320
 GACATTGCAG AAGCTTGAAA GGCCAATACC AGAACACAGG CTGATGCTTC TGAGAAAGTC 1380
 TTTTCTTAGT ATTTAACAGA ACCCAAGTGA ACAGAGGAGA AATGAGATTG CCAGAAAGTG 1440
 ATTAACCTTG GCGGTGCAA TCTGCTCAAA AACTGAAAC AATAAATCTG 1500
 ACCACTCCTA TGTTCGGACC CAAGCAAGTT AGCTAAACCA AACCAACTCC TCTGCTTGT 1560
 CCCTCAGGTG GAAAAGAGAG GTAGTTTAGA ACTCTGCA TAGGGGTGGG AATTAATCAA 1620
 AAACCKCAGA GGCTGAAATT CCTAATACCT TTCCTTTATC GTGGTTATAG TCAGCTCATT 1680
 TCCATTCCAC TATTTCCCAT AATGCTTCTG AGAGCCACTA ACTTGATTGA TAAAGATCCT 1740
 GCGCTCTGCT AGTGTAACCT ACAGTAAGTC TAAAGATGAR AGAGTTTAGG GACTACTCTG 1800
 TTTTAGCAAG ARATATTKTG GGGGTCTTT TGTTTTAACT ATTTGTCAGG GATTGGGCTA 1860
 RAGAGACGAC GACGAGAGTA AGGAAATAAA GGRATTGCC TCTGGCTAGA GAGTAAGTTA 1920
 GGTGTTAATA CCGTAGAAAT ATGTAAGGGA TAGACCTCC CTTTCTTTAT GTGCTCACTG 1980
 AGGATCTGAG GGGACCTCTG TAGGAGAGCA TAGCATCATG ATGTATTAGC TGTTCACTG 2040
 CTACTGTTTC GATGGACATA ACTATTGTAA CTATTCACTA TTTACTGGTA GGCATGTC 2100
 TCTGATTAAA CTGGCCCTAC TGCAATGGC TACTTAGGAT TGATCAAGG GCCAAAGTGC 2160
 AGGTGCGGTG AACTTTATTG TACTTTGGAT TTGTTAACC TGTTTCTC AAGCCTGAGG 2220
 TTTTATATAC AAATCCCTG AATACTCTTT TTGCTTGTG TCTTCTCAGC CTCCTAGCCA 2280
 AGTCCTATGT AATATGGAAC ACAACACTG CAGACTGTAG ATTCAGTTGC CGATCAAGGC 2340
 TCTGGCATTC AGAGAACCCT TGCAACTCGA GAAGCTGTTT TTAATTGCTT TTTGTTTGA 2400
 TCCAGTGCTC TCCATCTAAT CAACTAAACA GGAGCCATTI CAAGGCGGGA GATATTTTAA 2460
 ACACCCAAAA TGTGCGGTCT GATTTTCAA CTTTAAACT CACTACTGAT GATTCCTCAG 2520
 CTAGGCGAAT TTGTCCTAAC ACATAGTGTG TGTGTTTTGT ATACACTGTA TGACCCCAAC 2580
 CCAATCTTT GTATTGTCCA CATTCTCCAA CAATAAGACA CAGAGTGGAT TTAATTAAGC 2640
 ACACAAATGC TAGGGCAGAA TTTTGAGGGT GGGAGAGAAG AAAAGGGAAA GAAGCTGAAA 2700
 ATGTAAACC ACACAGGGA GGAATAATGA CATTGAGAAC CAGCAACAC TGAATTTCTC 2760
 TTGTTGTTTT AACTCTGCCA CAAGAATGCA ATTTGTTAA TGGAGATGAC TTAAGTTGGC 2820
 AGCAGTAATC TCTTTTATG AGCTTGTACC ACAGCTTTC ACATAAGTGC AGATTGCT 2880
 CAAGTAAAGA GAATTTCTCT AACACTAACT TCAGTGGAT AATCAGCAGC GTAACCTACC 2940
 TAAAGACATA TCACTAGCCA AAGAGGGAAT TATCTGTCT TCTTACTGTG CCTATATTAA 3000
 GACTAGTACA AATGTGGTGT GTCTTCCAAC TTTCATTGAA AATGCCATAT CTATACCATA 3060
 TTTTATTCGA GTCACTGATG ATGTAATGAT ATATTTTTC ATTATTATAG TAGAATATT 3120
 TTATGGCAAG ATATTGTGG TCTTGATCAT ACCTATTAAT ATAATGCCAA ACACCAATA 3180
 TGAATTTTAT GATGTACACT TTGTGCTTGG CATTAAAGA AAAAACAACA CATCTCGAA 3240
 GTCTGTAAGT TGTTTTTTGT TACTGTAGGT CTTCAAGTT AAGAGTGTA GTGAAAATC 3300
 TGGAGGAGAG GATATTTTCC ACTGTGTGGA ATGTGAATAG TTAATGAAA AGTTATGGTT 3360
 ATTTAATGTA ATTTACTT CAAATCCTT GGTCACTGTG ATTTCAAGCA TGTTTCTT 3420
 TTCTCTTTA TATGACTTTC TCTGAGTTG GCAAGAAGA AGCTGACACA CCGTATGTTG 3480
 TTAGAGTCTT TTATCTGGTC AGGGGAAACA AAATCTTGAC CCAGCTGAAC ATGCTCTCCT 3540
 GAGTCAGTGC CTGAATCTTT ATTTTAAAT TTGAATGTTT TTAAGAGTT AACATTTCTA 3600
 AAGCAATATT AAGAAGACT TTAATGTTA TTTTGAAGA CTACGATGC ATGTATACAA 3660
 ACGAATAGCA GATAATGATG ACTAGTTCAC ACATAAAGTC CTTTAAAGGA GAAATCTAA 3720
 AATGAAAAGT GGATAAACAG AACATTTATA AGTGATCAGT TAATGCCATA GAGTGAAAGT 3780
 AGTTCTATTG ACATCTCTCA AGATATTATA TATCAACTGC ATTATGTATT ATGTCTGCTT 3840
 AAATCATTTA AAAACGGCAA AGAATTATAT AGACTATGAG GTACCTTGCT GTGTAGGAGG 3900
 ATGAAAGGGG AGTTGATAGT CTATAAACAC TAATTTGGCT TCAAGTTTCA TGAATCTGTA 3960
 ATAGAATT TATTTTCAAC CCAATAATGT TCTATATAGC CTTTGTAAAG GAGCAACTAA 4020
 TAAATTAAC CTATCTTTC AAAAAAAA

Seq ID NO: 461 Protein sequence
 Protein Accession #: NP_037504.1

1 11 21 31 41 51
 MSRTAYTVGA LLLLLGLLLP AAEKKGKGSQ GAIPPPDKAQ HNDSEQTQSP QPQSRNRGR 60
 GQGRGTAMPG EKVLESSQEA LHVTERKYLK RDWCKTQPLK QTIHEEGCNS RTIINRFYCG 120
 QCNSEFYIPRH IRKEGSGFQS CSFCKPKKFT TMMVTLNCPB LQPPTRKKRV TRVKQCRGIS 180
 IDLD

Seq ID NO: 462 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..2733

1 11 21 31 41 51
 ATGAAAGTTG GAGTGCTGTG GCTCATTCT TTCTTCACCT TCACTGACGG CCACGGTGGC 60
 TTCCTGGGGA AAAATGATGG CATCAAAACA AAAAAAGAAC TCATTGTGAA TAAGAAAAAA 120
 CATCTAGGCC CAGTCGAAGA ATATCAGCTG CTGCTTCAGG TGACCTATAG AGATTCCAAG 180
 GAGAAAAGAG ATTTGAGAAA TTTTCTGAAG CTCTTGAAGC CTCCATTATT ATGGTCACAT 240
 GGGCTAATTA GAATATACAG AGCAAAAGSCT ACCACAGACT GCAACAGCCT GAATGGAGTC 300
 CTGCAGTGTA CCGTGTGAAGA CAGCTACACC TGGTTTCTCT CCTCATGCCT TGATCCCGAG 360
 AACTGTCTACC TTCAACGGC TGGAGCACTC CCAAGCTGTG AATGTCTCT CAACAACTC 420
 AGCCAGAGTG TCAATTCTCG TGAGAGAACA AAGATTGGG GCACCTTCAA AATTAATGAA 480
 AGGTTTACAA ATGACCTTTT GAATTCATCT TCTGCTATAT ACTCCAAATA TGCAAAATGA 540
 ATTGAATTC AACTTAAAA AGCATATGAA AGAATTCAG GTTTTGAAGT GGTTCAGGTC 600
 ACCCAATTC GAAATGGAAG CATGTTGCT GGGTATGAAG TTGTTGGCTC CAGCAGTGCA 660
 TCTGAAGTGC TGTGAGCCAT TGAACATGTT GCCGAGAAGG CTAAGACAGC CCTTCAAG 720
 CTGTTTCCAT TAGAAGACGG CTCTTTCAGA GTGTTTGGAA AAGCCAGTG TAATGACATT 780
 GTCTTTGGAT TTGGGTCCAA GGATGATGAA TATACCTGCT CCGCAGCAG TGGCTACAGG 840
 GGAACATCA CAGCAAGTG TGAGTCTCT GGGTGGCAGG TCATCAGGGA GACTTGTGTG 900
 CTCTCTCTGC TTGAAGAACT GAACAAGAT TTAGATGTA TTGTAGGCAA TGCCACTGAG 960
 GCAGCTGTGT CATCTTCTGT GCAAAATCTT TCTGTCTACA TTGCGCAAAA CCCATCAACC 1020

	ACAGTGGGGA	ATCTGGCTTC	GGTGGTGTG	ATTCTGAGCA	ATATTTCATC	TCGTCTACTG	1080
	GCCAGCCATT	TCAGGGTGT	CAATTCAACA	ATGGAGGATG	TCATCAGTAT	AGCTGACAAT	1140
	ATCCTTAATT	CAGCCTCAGT	AACCAACTGG	ACAGTCTTAC	TGCGGGAAGA	AAAGTATGCC	1200
5	AGCTCAGGGT	TACTAGAGAC	ATTAGAAAAC	ATCAGCACTC	TGGTGCCTCC	GACAGCTCTT	1260
	CCTCTGAATT	TTTCTCGGAA	ATTCAATTGAC	TGGAAGGGGA	TTCCAGTGAA	CAAAAGCCAA	1320
	CTCAAAAGGG	GTTACAGCTA	TCAGATTAAA	ATGTGTCCCC	AAAATACATC	TATTCCCATC	1380
	AGAGGCGGTG	TGTTAATTGG	GTCAGACCAA	TTCCAGAGAT	CCCTTCCAGA	AACTATTATC	1440
	AGCATGGCCT	CCTGACTCTC	GGGGAACATT	CTACCCGTTT	CCAAAAATGG	AAATGCTCAG	1500
10	GTCAATGGAC	CTGTGATATC	CACGGTTATT	CAAAACTATT	CCATAAATGA	AGTTTTCCTA	1560
	TTTTTTTCCA	AGATAGAGTC	AAACCTGAGC	CAGCCTCATT	GTGTGTTTGG	GGATTTCAGT	1620
	CATTTCGAGT	GGAACGATGC	AGGCTGCCAC	CTAGTGAATG	AAACTCAAGA	CATCGTGACG	1680
	TGCCAATGTA	CTCACTTGAC	CTCCTTCTCC	ATATTGATGT	CACCTTTTGT	CCCCCTTACA	1740
	ATCTTCCCCG	TTGTAATAATG	GATCACCTAT	GTGGGACTGG	GTATCTCCAT	TGGAAGTCTC	1800
15	ATTTTATGCC	TGATCATCGA	GGCTTTGTTT	TGGAAGCAGA	TTAAAAAAG	CCAAACCTCT	1860
	CACACAGTGC	GTATTTCGAT	GGTGAACATA	GCCTGTGCC	TCTTGATTGC	TGATGTCTGG	1920
	TTTATTGTTG	GTGCCACAGT	GGACACCACG	GTGAACCCCT	CTGGAGTCTG	CACAGCTGCT	1980
	GTGTCTCTTA	CACACTTCTT	CTACCTCTCT	TTGTTCTTCT	GGATGCTCAT	GCTTGGCATC	2040
	CTGCTGGCTT	ACCGGATCAT	CCTCGTGTTC	CATCACATGG	CCCAGCATTT	GATGATGGCT	2100
20	GTGGATTTT	GGGGTGGTTA	TGGGTGCCCT	CTCATTATAT	CTGTCAATAC	CATTGCTGTC	2160
	ACGCAACCTA	GCAATACCTA	CAAAAGGAAA	GATGTGTGTT	GGCTTAACCTG	GTCCAAATGA	2220
	AGCAAAACAC	TCTTGGCTTT	TGTTGTCCCT	GCACTGGCTA	TTGTGGCTGT	GAACCTTCCT	2280
	GTGGTGTGAC	TAGTTCTCAC	AAAGCTCTGG	AGGCGGACTG	TTGGGGAAAG	ACTGAGTGGG	2340
	GATGACAAGG	CCACCATCAT	CCGCGTGGGG	AAGAGCCTCC	TCATTCTGAC	CCCTCTGCTA	2400
25	GGGCTCACCT	GGGGCTTTGG	AATAGGAACA	ATAGTGGACA	GCCAGAATCT	GGCTTGGCAT	2460
	GTTATTTTGG	CTTTACTCAA	TGCATTCCAG	GGATTTTTAA	TCTTATGCTT	TGGAATACTC	2520
	TTGGACAGTA	AGCTGCGACA	ACTTCTGTTC	AACAAGTTGT	CTGCCCTAAG	TTCTTGGGAG	2580
	CAAAACAGAA	AGCAAACTC	ATCAGATTTA	TCTGCCAAAC	CCAAATTCCT	AAAGCCTTTC	2640
	AACCCACTGC	AAAACAAAGG	CCATTATGCA	TTTTCTCATA	CTGGAGATTC	CTCCGACAA	2700
30	ATCATGCTAA	CTCAGTTTGT	CTCAAATGAA	TAA			

Seq ID NO: 463 Protein sequence
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
35	MKVGVNLIS	FFTFDGHGG	FLGKNDGIKT	KKELIVNKKK	HLGPVEEYQL	LLQVTYRDSK	60
	EKRDLRNLFLK	LLKPLLWSH	GLIRIIRAKA	TTDCNSLNGV	LQCTCEDSYT	WFPSPCLDPQ	120
	NCYLHTAGAL	PSCECHLNNL	SQSVNFCERT	KWGTFFKINE	RPTNDLLNSS	SAIYSKYANG	180
	IEIQLKAYE	RIQGFESVQV	TQFRNGSIVA	GVEVVGSSSA	SELLSAIEHV	AEKARTALHK	240
40	LFPLEDGSFR	VFGKAQNDI	VFGFGSKDDE	YTLPCSSGYR	GNITAKCESS	GQWVIRETCV	300
	LSLLEELNKN	PSMIVGNATE	AAVSSFVNQL	SVIIRQNPST	TVGNLASVVS	ILSNISLSL	360
	ASHFRVSNST	MEDVISIADN	ILNSASVTNW	TVLLRECKYA	SSRLLETLEN	ISTLVPPAL	420
	PLNFSRKFID	WKGIPVNKQ	LKRGYSYQIK	MCPQNTSIPI	RGRVLIGSDQ	FORSLPETII	480
45	SMASLTGNI	LPVSKNGNAQ	VNGPVISTVI	QNSINEVFL	FFSKIESNLS	QPHCVFDFPS	540
	HLQWMDAGCH	LVNETQDIVT	CQCHLTSFS	ILMSFPVPST	IFPVVKVITY	VGLGISIGSL	600
	ILCLIIIEALF	NKQIKKSQTS	HTRRICMVNI	ALSLIADVH	FIVGATVDTT	VNPSGCVCTAA	660
	VFPTHFFYLS	LFPMMLMGI	LLAYRIILVP	HMAQHLMMA	VGFCLGYGCP	LIISVITIAV	720
	TQPSNTYKRR	DVCHLWNSG	SKPLLAFFVP	ALAIIVAVNFV	VVLLVLTCLW	RPTVGERLSR	780
50	DDKATIRVIG	KSLILITPLL	GLTWGFGIGT	IVDSQNLAMH	VIPALLNAFQ	GFPIILCFGIL	840
	LDSKLRQLLF	NKLSALSSWK	QTEKQNSDL	SAKPKFKPFP	NPLQNKHYA	FSHTGDSNDN	900
	IMLTQFVSNE						

Seq ID NO: 464 DNA sequence
Nucleic Acid Accession #: AB035089.1
Coding sequence: 9845..10219

	1	11	21	31	41	51	
60	GGGCATGCAG	CCATCGGGGA	AAATCCATAG	TGCAGATAAA	GCAAGGAGGA	AGAAGAAGGA	60
	CAGTTCTAGT	AAAAGGGAGA	ACATCAATAT	AGGATGTTTC	TAGCAATAG	AAAAGAAGG	120
	CCAAGAGGAA	TTAGGGAGAG	AGTTATAAGA	GATCAGCAAG	GGGACAGGGT	TAGATTGGT	180
	TTGGTTTGAA	AGCATACAGT	AAATATGATG	TCTGTCCCTG	GCACTGTGTTG	CAGAGTAGGA	240
	AGGAGGAAGG	GAGGCAAGAG	ATAATATCAT	TTTCTCTGTG	CTCCAACCTG	ACTTACATAT	300
65	GAGACTATTT	CCCTCTCTGC	TTTTCAAACC	TTACTGGAGT	TGTTTTCCCT	CATGAAAACC	360
	AAGAAAGGAA	AGCTAGTTAG	TCTTGTCTG	AGGTTGTTC	ATGTATACAT	ATCTATATCT	420
	GTAGACAGAA	TCTTGGGAA	TACAGTAATT	GACATATATT	CTGTTATTTG	ATGCTTGAAA	480
	AATCTCCTCC	ACTAACCACT	TTCCCTATAG	ATTGCCACAA	GCACATAATA	AGAAACAATA	540
	AATAAAATGT	TCTCTTGACT	TTGTACTTA	ACAATGCTGA	GAAAACCTTA	CAGCCTTCAT	600
70	AAGGAAGTGA	GGTCCAGGAA	AATCTAGGAG	ATATTCTTA	ACCAATCTAT	AAAGGCATTA	660
	GTAATGACAG	GATATTTCTT	GAAAGTGTA	TTTCCCATG	AGGATTGTTT	TTTAATTTCT	720
	GGATTCTCTG	AGCCAAATGA	GTTGGTGTAT	GTTTATGAAA	TATCAAGAGA	CATAAGTTGG	780
	CAAGTGTTCA	TATGCAAAAA	CTTCTGGAA	TTTCTGAGTT	CTCTGTGGCA	ATATATGACA	840
	TCAGGATATG	TCCAGTCTCA	CACACCAGGA	TATGTCCTTT	CTAGCCTGTC	TATCACATGC	900
75	TAGGAGAACT	ATTTAGGAAC	AGAAAAAAT	GCCTGAAATG	ATTTCTCATT	TGAACCTATC	960
	CAAGCTTTCT	CYAAATTTAA	GCAAACTCCT	GGTCATTTTC	AGTTAGTACC	TTTCTTTAAG	1020
	TTCAACCTTC	AGSGCAAAAC	TCCGTGCCCT	AGAGGTTTAC	CCATAGTCTG	AAATCTCTCT	1080
	CCATAGATTG	GTCCCTCTGA	ACCCCGGTTT	GTCTCAGCTT	GTTATCCTGT	TTTTTCTTCT	1140
	CCTCCATTCC	CAGGATGAGC	TTGTGTCTTC	TGTCCTATGA	GACATTAGAT	TCCTTTTCTT	1200
80	TGGTACCCGA	GTAATCCAT	CCTACTCCAA	TAGAGGAAGS	TCCATTTTGT	TCTTATAGCG	1260
	CTGGATGCAG	ACTCAGCTGA	GAAGACCATT	ATTCAATTTT	GGAATTCCTT	ATCTCAGATA	1320
	TTTCTCTTCT	TTTCTTTTCT	TTCTATCTTT	GGATTTTATG	TCCATCAACG	CCCCATTAGT	1380
	CTATTCCCGG	ACTTCAATCA	GGGAACCTAT	ACCTCTTAAA	CTCATTGAGA	GACTCAAAAC	1440
	ATATATATTG	ATACAGGAGA	CCTAAGAGA	GCATGTCTTG	GGGGTTGAGG	AAACAGGCAG	1500
85	GTGAGAAATT	TCCAGATTGG	AAACACAGCT	TCCTTTCTCC	CATCCAGCCC	CTACTTTCAG	1560
	CCTATGTGTT	TCTGGCACCT	TGTTGTAGAT	AAATCTCCCT	TGACTTTGTG	ATGTGCTGAG	1620
	AAAACAAACT	CACGGCTGGT	GTTAAAAAGG	GCCCATGACA	ATACCAAGTG	TTGGGGAGAA	1680
	TGTGGAGAAA	TCAGAACTCT	ATTCAACGGT	GTTTGAATG	CACACTTGTG	CAGAATTCTA	1740

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85

TGGAGAAGAG TCTGGCATT TCTCAAATG TTAACCTGGA TTTACCATAT GACCCAGCGA 1800
 TTTCATTTCAT AGGTTTATAC TCAAAAGAAA TGAAGAAATA TGCCATGCAA AAAAATGTAC 1850
 ATGAAGAGTC ACAACATCAT TATTCATAAT AGTAAAAGGA TGGAAACAAC ACAAAATGTC 1920
 ATCAACTTAT GATTAAAGAA AATCTGGTCT ATTCTAGAAA TGGAAATATTA TTCGACCACA 1980
 AAAAGGAATG ATGTACTGAT CCATGCAATG ATGTGGACAA ACCATGAAAA TAACACTAGA 2040
 TTAAGAAAGC CAGTCACAAA AGGACTTACT GTATGATTCC ATTTACCTGA AATGTTTGGG 2100
 ATAGGCAAAAT CCATAGAAAAG AGGAGGTAGA TTCTCTGGTTT CCAGGGTCTC CAGGAAGGGA 2160
 AGAATGAAGT ACAAGATTTC TTTTGGAGGT AGTGAATTTG TTGTGGAAATG AGATCATGAT 2220
 GATGATAGCA CAACCTTTG TAAATAATAA AATCATTGAA TTGTACAGTT GAATTTATGG 2280
 TATATAAATT ATATGTTAAT AAAAAGGGGG TCCACAAAAC AAACAGCCCC CCACTCTGGT 2340
 TGTCAAGGAG ATATTGGATT AAATGGCCTT GGACAACAAC CCTCTCCCTT GGGCAGAGAC 2400
 ATTTCTCAGA TTACAAGATA TTCCAGGGGA AACACTGGAA TGAGTCTGAA GCCAGGTGCT 2460
 AAACAGAAAG ACCATTGAGA AATGTTGTGA TCCTGACAGG TCAAGCAATT TATTTTTCGG 2520
 CTTCATTTTT AAATGTAATA TTAGAAAGCT GCCATTAAA ATGGCCCGTC TGTTCATATT 2580
 GCTCTTCTCA GTGTGAGCCT GTTAACCTCAA TGTGTTAGTC TGTTCATG CTGCTGATA 2640
 AAACATACCT GAGACTGGCA AGAAAAGAG GTTTAATTGG GCTTAGAGTT CCACGTGATT 2700
 GGGGAGGCTT CAGAAATCACA GTAGGAGGCA AAAGTTATTC TTACATGGTG GCTGCAAGAG 2760
 AAGATGAAGT AGAAGCAAAA GAAGAAACCC CTGATAAACC CATCGGATCT CCTGAGGCTT 2820
 ATTAACATATC ATGAGAATAG CACAAGAAAG ACCGGCCCCC ATGATTCAAT TACCTCTACC 2880
 TGGTCCCCTC CAATAACATG TGGAAATTTT GGTAGATACA ATTCAGTTG AGATTGGGT 2940
 GGGAAACAGG CCAAAACATA TCACCTCAGCA AGGCAGATAA CTCTCTCACT GAGCCTATGC 3000
 AACAGAAAAC CATCTGGGAT GGTGTGAAGG GGCACAGGAA GTGACTGGTA GGATCACTGC 3060
 CAAAGCTGAG CACTCAGGAG AAGGCAATAG AATCCTATT TCATAGTAT GCTATAAGAT 3120
 ACTGAAGTAC ACTTCTTAC TATCTCTTG GACTTAGAAT TAGCACTACA TTCTTGTGTA 3180
 TACAGAAAAA TTACTAAGGA AATTCATAGG ATGACAAAAA CTTCAGAAC TGAAAAACAG 3240
 GAAATGTAAAG CTCTTTAGTT CTCTGTGATT CGAAGTATGC CTAAAAGACA ATGCAAAATC 3300
 CAAGAAAAGA ATGGTGGGT TTTTGTGTTG TTGGTTTGT TTTTGTGTTA CAGCTGGAGT 3360
 AGAATACAAA GGGATGGAGT TGAAACAAAT GAGAGGAAAT TGGAAATCTA AACTTATTCT 3420
 CATTTGGCATT AGAAAGGCAC CTACATGAT TTCACATGAG CCGTGACTG CTGACTTGCA 3480
 TTCTTATTTT TTCCCTATAG ATTAAGAAAG AGGTACAATG GTAGAATCTG AATCTGTGCC 3540
 TTTGTCAATA ATTTTCATAT TCATAAAGGT GAGTGTATGC CCGCTTGTGA AATCTGAAGT 3600
 TGAGTACATT CAAATACTAA CCACAGAGGG AAAGGCAGCA AGAGGAGAGG CATAAATTTA 3660
 GGAATCTACC CTTCATTCCA CAGACACACA CAGCCTCTCT GCCCACTCT CTCTCTCTA 3720
 GGAACACAGG TAAGAGCTTC AAGCCTCTCC AGCTTAATAA CATGAATTAT TTTTGAGAA 3780
 AATAATGATA CTGTGTTCTA TATCATGCAT CTCTCTGATT CTGTCTGATT ATATTTTACT 3840
 TATCTGCGCA GAGCAAAAT AAAATACCTA TTTCTATCTA TTGTCTCTT ATCTAAATG 3900
 CTGTAGTTCA AGTAAACCAA GGCACCTTTA GGAACACAGA GGGAGAGTGC CTGCGACCA 3960
 GAGAGTCTTG AAGGAGATGT CAGGACGCA TCTTAACAGC TGGTGGATG TGATCCACAG 4020
 AGGTCTCTCTG TTAGCATTC TTTGTAAGCC ATCTACCTA GCTCTAGTGT AACCCAGCAAT 4080
 GAAAGAAAGA TAAAGAGGTG CGATTACTTA TTTACAATAG TCTTTAAAAA CGTAGTTTG 4140
 TAAGCCTCTT AATTAGGACA TTAATATATT TAATATATGC ACATTGTAGA AAGATTGAAG 4200
 CGTTAAAAAT AAGAGAAAAA CTTTAAATGT CAAATCTCA CAACCCAGAT ATATCATTTT 4260
 TTTAAGAAAA TGTACTACA AATACCAT CCAATTATTA AAGTCAATCT GACAGGAATC 4320
 TGATGCTTTT CCAGGAGTTC CAGATCACAT CGAGTTTACC ATGAATTCAC TCAGTGAAGC 4380
 CAACACCAAG TTCATGTTG ATCTGTTCCA ACAGTTTACA AAATCAAAG AGAACACAT 4440
 CTCTTATTC CCTATCAGCA TCACATCAGC ATTAGGATG GTCTCTTAG GAGCCAAAGA 4500
 CAACACTGCA CACAAATTA GCAAGGTAGC TATCAGCATC ATTACGTTGT CCTGTGCGAG 4560
 TTTTCTCTG GTTCCCTCGG CTAGCAGCA GATGGTAATA GATGTGGTGG TCTGATGGGT 4620
 AGCACAGGGG GCTGTGCGAG AATTCACATA ACTGTGAGAC CACTGACTTA AACAGATCTT 4680
 TTGAGTAAAG TTTTCTTGT CCGCTTCATG TCTCTTCCAG GTTCTTCACT TTGATCAAGT 4740
 CACAGAGAAC ACCACAGAAA AAGCTGCAAC ATATCATGTG AGTCACAGAG CACTCTGATT 4800
 CAGCTTTAGA TCCTGAAACA GGTCTAGATT TAAACCTGGA ACTTCACAAA AACTAAGAAA 4860
 AGGCCATTAT TAGGAAAAAT CTGGACACA AAGATTGAGA CACACAGAGT GGGTTGGCAT 4920
 TTCTATGGCACA ATAATTATTA TTCTCTATT CTGCGTTACT AAAAGACAGT CAGCACTGTA 4980
 CCTCAGAGCA TAGGTCTGGA TCAGGATAGG CTGGGTTTCC ACTCCAGCTT TGCTCTTCC 5040
 AATGATGAAA TAAGAGCAGG ACACAACCTG TCGGAGTCCC AGTGACCTCA TCCCAGAAAA 5100
 CTAAGGGTAA GAAAAAATCT GACTCAATAC ATGCAATATC ATGCAATATG TTACAACAGT 5160
 GCCTTGCCCA TAAAGTCTAT AATAAATGTT ATTATTATTA TAAAGTAGCT ATAATTATAC 5220
 TAATCATAAAT AATGTGAAAA TAATTTAATT TTCATTGAGT CATTATAGAG ATTCAAGGGA 5280
 ATAAGCACAA GTCCAAGTAT ATTTTGGAAA ATGATTGCTA TGAATATAT TGGTTTAGAG 5340
 CCTTAATAGT GCAAAATGCT TTGCTGGAAG GTAGAAAGTT CTAGATTAA ACAGGCTTAG 5400
 GTTCAAACT TGGCACTTCT AATTTATGTC TCTATAACA GGGTTTTTTT CCCCATTCTC 5460
 TGAGCTTTCT TGTGTTTCTC TGAATTGAAC TAAAGACTTA GAGTTACCCA TGTAAAGTCC 5520
 TTAGCCATGG ACCTGGCATA CACTCTCTT ACGTGACAG AATGACCATC ATGAGGAAAG 5580
 AGCCACAGAT CAGTCAATGT GTCTTACAAG ATAATAGCAC CAACAGGTAT AACAGGGCTT 5640
 CCTGGCATAA TCTATTAAAA ATATCCAACC TTCAACATAC TCGTATCCTT GATGACTGTT 5700
 AGAAGTAAAA TATGTCCTT GCCCATAAGG AGCTGAGAGT TTAAGTGGGA AGCTAAACCT 5760
 AACCTTTTAA ACCAACAAAG AGAAAACTA CTGGTAGACA GCGCTGCATC TTTAGTTTCA 5820
 AAGAGAAAAA ATTGCAGTAC GTTAGAGCAA GAAGAATTTT CTGGAAGAAG TCAATATATA 5880
 GGTGGATTTT GAAGGGTATT TGAGGTGAAA TACACCAATT ATCAGGGAAT AACATCAAAG 5940
 GTCTCAATG AGACTACCA CATTAGGGA CTGATCTAAC AGACTTAGCA TGGGTTTAGT 6000
 ATTTACATTG ATACAGCAAT TGAATGATCT CCTTTTGTGA TGTGTAAGG TTGATAGGTC 6060
 AGGAAATGTT CACTACAGT TTCAAAGCT TCTGACTGAA TTCAACAAAT CCACTGATGC 6120
 ATATGAGCTG AAGATCGCCA ACAAGCTCT CCGAGAAAAG AGTATCAAT TTTTACAGGT 6180
 AATTTCACTT GGCCTACCCA CATTTCATT GCATCTGAT GTCTGTGCT CTGAGTGGCC 6240
 AATGGAAGA AAGCAAGGCA GATGAGCCTG GCGACCCAG GTGGAGAGCA TTTACTCAGA 6300
 GTGCATTAGC TCCATTCCA CACTCTCCC CCACTGGAGT GTCCAGAGCC CCAACGATAC 6360
 ATCACTGAAG TGTGGATTTA GGGATAATCT TGTGATAAAA GAGGAGGTTG TGTAAATAGAG 6420
 TGAGTAAGAG TAATAAGTAA TAAGATACCA TCGATAAACT GGCCTGACT CAGTCACATA 6480
 CGATACATCT TGGTGGGAAA TGTATGACTA ATGGGATATT ATTGGAATGG GCAGGCTTGG 6540
 GTGATTCCT GAGAAATGTT GAGGAAGTAC CAGGAATAT TGAATGCACA GGATGAAAGA 6600
 CAAAACAAA GATCAGAAAC ATCATGGTTA AAATTACTGG AGAGAAGTCT GAGAAGCAAT 6660
 GAATCTCCTT CAGGGAAGCC TGCTCTGAG TTTGCAACC ACAGCCTCT CTGCTCTGCT 6720
 CTTTTGCCAA GATGATATTG ACCTTCAGT ACCTCTTCT TGTGCCAGCC CACATTCCCC 6780
 TTTTGATTG CCTACATGAC ACCTGTATAA AAATATCCAT GGACAGGAGA TACTGCATCT 6840
 ATTCAGGCTC TGGATTGAG TTACTGTTGT TACAATAAG TAAGTTTGGT AATATATAGT 6900
 TACATAAATT ACTCTAATT CTTACTTCTT CTTTCATATC TCAAAGGAAT ATTTAGATGC 6960

5	CATCAAGAAA	TTTTACCAGA	CCAGTGTGGA	ATCTACTGAT	TTTGCAAATG	CTCCAGAAGA	7020
	AAGTCGAAAG	AAGATTAACT	CCTGGGTGGA	AAGTCAAACG	AATGTAGGA	GAGCCACCCA	7080
	TTATAGAAAC	ACCTTTGAGA	AACCTATGCC	AGTGAGCCTT	GTGCTTGACA	CTGCATGGGG	7140
	GAAACAGTGT	GGGGATTGAG	ATGGGTTTGC	AGGGAGGGCT	GAAGAGGGCA	CTCCAGATGA	7200
	AGGATTGTGC	CAAAATGAAT	TGAAGAGAGC	CTAGGGGAGC	CAAGGAGGAA	ATCACAGGAA	7260
	GCCAATTAGA	TGGAAACACA	TCTGGAGAAT	TATTTGCTTA	TGGCCCTGCA	TGACAATAGC	7320
	TTTGTGGATC	CCCTGTCTCC	GCTCAGACCT	ATTTTGAGAT	CATATCCTTT	ACTTTAAATC	7380
	AGACTCAAAT	TTTATGATG	AATATTTAAT	AGAAAACTT	AGAAAGCCTC	TCTCGTCTCC	7440
10	TTTACTAATT	GGGAAACAAG	CAGCTCTCTG	GTAATACACC	CTTTTGTCTC	TGAGCTGGAG	7500
	CTGCCTGGAT	CACATCTGTA	GCCAAATGTG	TCTGCAGGGA	TTATCACAGC	TCTCTTCCCC	7560
	ATCAAGGGCA	AAGAGCTTGA	CAAAGTCTCC	ATTCTACAGA	CATCTTCTT	ACCTCCCAAC	7620
	TCTCATTACA	GGCCAAACTT	ACAGCAACTC	AACATGAGAG	TGAATAGGAA	GATACCCCGG	7680
	GAAGTAGTGT	CTGACAGCAC	AGGACATGGG	TTTCATATTA	CAGAGCTCAA	GTCACATC	7740
15	CTAAAATGCA	ATCAGGGCCT	CCTTCTCTG	AATGGGGACC	CGTAGTTAA	AAAAAATAA	7800
	AAGTAGGAAG	AGGAGGGAGG	GAGAAAGGAA	AGACACATGT	TGGAAGAGTA	GACAAAATCA	7860
	GTTTATCAGT	ATTCCAAATC	AGATGATTGG	AGACATTCAT	ACACAGAGAA	CGTGAACCTC	7920
	TTCTCTATCA	CAAGAAGTGA	TGTCCTCATC	AAGGGTAACT	TTATACGACT	GGAGCCTTGA	7980
	AGAAAGCTGC	ATCTGGTGAA	CCACTGGTCA	GTGAGTCTAA	CAATTCAAAG	ATCAAAGTCA	8040
20	GTGAGTCTCA	AGCAGGGATT	TGGGTCAATA	ATTAACGATC	AGTCACGAAC	ATTGCAAAAG	8100
	CATCTCTAAG	ACAGGCTATT	TGTAGCTTGT	GTAAAGAGCT	CTTTTATCTT	TCCCTTGCCA	8160
	GAAAAAATTA	AAACCTATT	TCTGATGGG	ACTATTGGCA	ATGATACGAC	ACTGGTTCTT	8220
	GTGAACGCAA	TCTATTTCAA	AGGGCAGTGG	GAGAATAAAT	TTAAAAAGA	AAACACTAAA	8280
	GAGGAAAAAT	TTGGCCAAA	CAAGGTATTG	TCTATATTTT	ATTATATAG	TGTAATATGT	8340
25	TAATACATGG	AATGTTAAAC	ATTTCTGATG	GAATGTAACA	TGATAAGTAA	AAAAAATAAA	8400
	TGTTCTATGT	CTGTTATTTT	GTGTTTATAC	TCTTATAACT	TTATTTAGTT	AGGAATACCT	8460
	GAAAAAATAT	TGTTTCTAAC	TCATGGAATT	CCTGGGTAT	TTCTTAGAAG	AAGAAGGATG	8520
	TGTTGCTATC	TCAATAATAT	TATCTTTTTT	GTCTTGTGTT	TCACGTGTTA	TTTGTGGAC	8580
	ACATTGATTT	ATTGCAGAA	ACATACAAAT	CTGTACAGAT	GATGAGGCAA	TACAATTCCT	8640
30	TTAATTTTGC	CTTGCTGGAG	GATGTACAGG	CCAAGGTCTC	GGAAATACCA	TACAAGGCA	8700
	AAGATCTAAG	CTGATGTTGT	CTGCTGCCAA	ATGAAATCGA	TGGTCTGCAG	AAGGTAAAGAA	8760
	CTTGCACTCA	CAACTCTTCC	TTCTACTGCC	GGACATTTTT	CCAAAGATAC	CAAGTTTAAA	8820
	CAAGGTAAAA	GCTTATGACC	GAGTTGCCTC	AAAAATGATGA	AAAATCTTAA	ATGAGGAATG	8880
	ATGACTCAC	TTTATATTAC	AAATATTGGA	GCATAGGGCC	TGACACAAAC	TGAAAGCTTA	8940
35	GTGTTTGT	GTGTTTGT	TTTTATTATT	ATTATTATAA	TACTTTAAGT	TTTGGGTAC	9000
	ATGTGCACAA	TGTGCAGGTT	AGTTACATAT	GTATACATGT	GCCATGCTGG	TGTGCTGCAC	9060
	CCATTAACTC	ATCATTTAGC	GTTAGGTATA	TCTCCTAATG	CTATCCCTCC	CCCCCCCCC	9120
	CACCCACCAA	CAGTCCTCAG	AGTGTGATGT	TACCTTCTGT	TGTCCAAGTG	TCTCATTTGT	9180
	TCAATTCCCA	TCTATGATTT	AATTCCTATC	ATGGCTTAGT	TAATGATTAA	TTTATTAGAG	9240
40	TTACATGCAT	TGGATATCAA	TTTATGATA	TTATATTGCA	GCAATTTAAA	CTTGACTGGG	9300
	AGAAATATAT	CTCAATGTGA	GGAAAGTTTA	CAAATAGGCC	GAGTAGAAAA	GGGAATACAA	9360
	ATTTAGGAAT	TTAGGGAATT	ACAATTTAAT	AATTGCAATG	TGTACTAAAT	AATGTATACA	9420
	GAAAAATATG	ATGAGCCTAT	TAAAAATTGA	CACATGTAGT	AGGCTGTTGG	CACAAGAAAT	9480
	AGTGATACAT	ACAGTTCAAT	GTGTACAAAA	TAATGTAATC	ATATTTTACA	TGTGTATCAT	9540
45	ACAGTTGTAT	ACATACATAT	GTACACATAT	ACATATACGT	AAAAACATGA	TTCTGTTTTT	9600
	ACATACATGT	ATATACATAT	ACACATATAA	CCCAATGTAT	TTATATATTC	AGGACTCATA	9660
	TTTTACCTAT	TAGAATAATA	ATGCTATTA	AGTGAAACCT	TCTGTATTTC	ACATTATTGT	9720
	CCAAAATAAC	GAATCTCCAC	ATAGTCAATT	CATTGTTAAG	GTGTATTAGA	GATCGACAGT	9780
50	TAGTCATATC	AGTTTCTTTT	TTCCATTGT	ATAGCTTGAA	GAGAACTCA	CTGCTGAGAA	9840
	ATTGATGGAA	TGGACAAGTT	TGCAGAAAT	GAGAGAGACA	TGTGTCGATT	TACACTTACC	9900
	TGGGTTCAA	ATGGAAGAGA	GCTATGACCT	CAAGGACACG	TTGAGAAACN	TGGGAATGGT	9960
	GAATATCTTC	AATGGGGATG	CAGACCTCTC	AGGCATGACC	TGGAGCCACG	GTCTCTCAGT	10020
	ATCTAAAGTC	CTACACAAGG	CCTTTGTGGA	GGTCACTGAG	GAGGGAGTGG	AAGCTGCAGC	10080
	TGCCACCGCT	GTAGTAGTAG	TGGAATTATC	ATCTCCTTCA	ACTAATGAAG	AGTCTGTGTG	10140
55	TAATCACCTT	TTCTTATCT	TCATAAGGCA	AAATAAGACC	AACAGCATCC	TCTTCTATGG	10200
	CAGATTCTCA	TCCCATAGA	TGCAATTAGT	CTGTCACTCC	ATTTAGAAAA	TGTTACCTTA	10260
	GAGGTGTCT	GGTAACTGA	TTGCTGGCAA	CAACAGATTC	TCTTGGCTCA	TATTTCTTTT	10320
	CTATCTCATC	TTGATGATGA	TAGTCATCAT	CAAGAATTGA	ATGATTAAAA	TAGCATGCCT	10380
	TTCTCTCTTT	CTCTTAATAA	GCCACATAT	AAATGTACTT	TTCTTCCAG	AAAAATTTC	10440
60	CTTGAGGAAA	AATGTCCAA	ATAAGATGAA	TCATTTAATA	CCGTGTCTTC	TAAATTTGAA	10500
	ATATAATTCT	GTGTTCTGACC	TGTTTTAAAT	GAACCAAAAC	AAATCATACT	TTCTCTTCAA	10560
	ATTTAGCAAC	CTAGAAACAC	ACATTTCTTT	GAATTATAGT	GATACCTAAA	TCCTTCTTAT	10620
	GTGTTCAAAT	TTTGTGATTC	TATAAAACAC	ATCATCAATA	AAATAATGAC	ATAAAATCAT	10680
	TTTTGCTTTA	CCTGTTTTCT	CTCTGGAAG	GGCAAGTGTG	CAGTTACACA	TAGGAAGAT	10740
65	AATTTAGAGA	TATATTAATC	ATATATAAAG	GAAATTTAAA	AACAGAGTAG	TTCATGATGA	10800
	GCCTGGAGTA	GAAGGCATAT	CCGAGAACAG	GAGGAGCCTT	GTAACCCACA	TAGGAACCTC	10860
	CTATTTTATG	CTAAAGGGAT	AAGAACTCA	TTACAGGCTT	TGATGGTTGT	TTGTCAAAGA	10920
	GGGGCATAAA	ATTATCATAT	CCACATCTAG	AAAATACATC	TCTGGCTACG	CTGATATCAA	10980
	TGGATGCGAG	GAAAGAACAG	TGTGGTTACC	ATATATAAAT	TAGGAATCA	TTAGAGTATT	11040
70	GGGAGTGGAA	ATGGAGAGAA	AGAAAGAGCC	TGGGGGAATT	ATTTAGGAAA	TAATAGTTAC	11100
	AGAAAGACAT	CTAAGTTGCT	GACCTATCTG	ACTGGATGGA	TGGAAGAATA	TCTTGTCTCT	11160
	GAGAGAAAAA	AAGACTTTGG	GTTTAAATTT	GTACTTGATG	AATTAAGGTA	CTTTAATAT	11220
	TCAATGGGAT	TTGCTGGCAA	GGCACTTGAA	GATATTAGTC	TAAATCTCAG	AAACAGAATA	11280
	TGATCTGAAG	CTCTAAATTT	GTGATATTCA	ATATAAATAC	TTTAGAGTCA	TTGGGATAAA	11340
75	TATGGTAGTT	GTAGCTAAAA	GCAAAAAATA	GATACTAGGG	AGAAAGGATA	AAGTTAGAG	11400
	AAAGAAGAA	CTGAATTTGA	CCTTGAAGTA	TATCAGCATG	TGTAAGATC	AGGAATTGAT	11460
	CATTTTTATT	TTCCAGAAAG	TAGCTTTTCT	TAGGGTTCCA	TATTTACTCC	CATAGATTCT	11520
	TCCC						

Seq ID NO: 465 Protein sequence
Protein Accession #: BAB21525.1

1	11	21	31	41	51	
MNSLSEANTK	FMFDLFQQFR	KSKENNIFYS	PISITSLGM	VLLGAKDNTA	QQISKVLHFD	60
QVIENTTEKA	ATYHVDKSGN	VHQFOKLLT	EFNKSTDAYE	LKIANKLPG	KTYQLQEYL	120
DAIKKPYQTS	VESTDFANAP	EESRKKINSW	VESQTNKIK	NLFPDGTIGN	DTTLVLVNAI	180
YFKGQWENKF	KKENTKEEKF	WPNKNTYKSV	QMMRQYNSFN	FALLEDVQAK	VLEIPYKGGD	240

LSMIVLLPNE IDGLQKLEEK LTAELMEWT SIQNMRETCV DLHLPRFQME ESYDLKDTLR 300
TMGMVNIENG DADLSGMTWS HGLSVSKVLH KAFVEVTEEG VEAATAAVV VVELSSPSTN 360
EEPCNHPFL FFIRQNTNS ILFYGRFSSP

Seq ID NO: 466 DNA sequence
Nucleic Acid Accession #: NM_001910.1
Coding sequence: 50..1240

1 11 21 31 41 51
GGAGAGAAGA AAGGAGGGGG CAAGGGAGAA GCTGCTGTGC GGACTCACA TGAAAAAGCT 60
CCTTCTTTTG CTGCTGGTGC TCCTGGAGCT GGGAGAGGCC CAAGGATCCC TTCACAGGCT 120
GCCCTCAGG AGGCATCCGT CCTCAAGAA GAAGCTGCGG GCACGGAGCC AGCTCTCTGA 180
GTTCTGGAAG TCCATAATT TGGACATGAT CCAGTTCACC GAGTCTGTCT CAATGGACCA 240
GAGTGCCAAG GAACCCCTCA TCACTACTTT GGATATGGAA TACTTCGGCA CTATCTCCAT 300
TGGCTCCCCA CCACAGAACT TCACTGTCTT CTTCGACACT GGCTCTCTCA ACCTCTGGGT 360
CCCTCTGTG TACTGCACTA GCCAGCCTG CAAGACGCAC AGCAGGTTC AGCCTTCCCA 420
GTCCAGCACA TACGGCCAGC CAGGTCAATC TTTCTCCATT CAGTATGGAA CCGGGAGCTT 480
GTCCGGGATC ATTGAGCCCG ACCAAGTCTC TGTGGAAGGA CTAAACCGTG TTGGCCAGCA 540
GTTTGGAGAA AGTGTACAG AGCCAGGCCA GACCTTTGTG GATGCAGAGT TTGATGGAAT 600
TCTGGGCTGT GGATACCCCTT CCTTGGCTGT GGGAGGAGTG ACTCCAGTAT TTGACAACAT 660
GATGGCTCAG AACCTGGTGG ACTTGGCGAT GTTTTCTGTC TACATGAGCA GTAACCCAGA 720
AGGTGGTGGC GGGAGCGAGC TGATTTTGG AGGCTACGAC CACTCCCAT TCTCTGGGAG 780
CCTGAATTGG GTCCAGTCA CCAAGCAAGC TTAAGTGGAG ATTGCACTGG ATAAACATCA 840
GGTGGGAGGC ACTGTTATGT TCTGCTCCGA GGGCTGCCAG GCCATTGTGG ACACAGGGAC 900
TTCCTCATC ACTGCCCTT CCGACAAGAT TAAGCAGCTG CAAAACGCCA TTGGGCGAGC 960
CCCGCTGGAT GGAGAAATAG CTGTGGAGTG TGCCAACTT AACGTCAATG CCGATGTCTC 1020
CTTCACTAT AACGGAGTCC CCTATACCT CAGCCCAACT GCCTACACCC TACTGGACTT 1080
CGTGGATGGA ATGCAGTCTT GCAGCAGTGG CTTCAGAGGA CTGACATCC ACCCTCCAGC 1140
TGGGCCCCCT TGGATCTCTG GGGATGCTCT CATTCGACAG TTTTACTCAG TCTTTGACCG 1200
TGGGAATAAC CGTGTGGGAC TGGCCCCAGC AGTCCCTTAA GGAGGGGCTT TGTGTCTGTG 1260
CTGCTGCTGT TGACAGACCT TGAATATGTT AGGCTGGGGC ATTCTTTTACA CCTACAAAA 1320
GTTATTTTCC AGAGAAATGA GCTGTTTCCA GGGTTGCAAC TTGAATTAAAG ACCAAACAGA 1380
ACATGAGAAT ACACACACAC ACACACATAT ACACACACAC ACATTCACA CATAACACC 1440
ACTCCACCA CGTCTATGAT GGAGGAATTA CGTTATACAT TCATATTTTG TATTGATTTT 1500
TGATATGAA AATCAAAAAA TTTACATTTT GATTATGAAA ATCTCCAAC ATATGCACAA 1560
GCAGAGATCA TGGTATAATA AATCCCTTTG CAATCCACT CAGCCCTGAC AACCCATCCA 1620
CACAGGCCCA GGCCTGTTTA TCTACACTGC TGCCCACTCC TCTCTCCAGC TCCACATGCT 1680
GTACCTGGAT CATTCTGAAG CAAATCCGA GCATACATC ATTTTGTCCA TAAATATTTC 1740
TAACATCCTT AAATATACAA TCGGAATTC AGCATCTCC ATGTGCCAC AAATGTTTGG 1800
CTGTTTTTGT AGTTGGATTG TTTGTAATTG GATTCAAGCA AGGCCATAT ATTGCATTTA 1860
TTTGAATGT CTGTAAGTCT CTTCATCTCT ACAGAGTTTA GCACATTGTA ACGTTGCTGG 1920
TTGAAATCCC GAGGTGTCTA TTGACATGGT TCTCTGAAT TATCTTTCTT ATAAATGGT 1980
AGTTAGATCT GGAGGTCTGA TTTTGTGGCA AAAATACTTC CTAGGTGGTG CTGGGTACTT 2040
CTTGTTCAT CCTGTACAGA GGCAGATAAT GCTGGTGCCT CTCTATTGGT AATGTTAAGA 2100
CTGCTGGGTG GGTGGGAGT TCTGGCTTT AATCATTCAT TACAAAGTTC AGCATTTT

Seq ID NO: 467 Protein sequence
Protein Accession #: NP_001901.1

1 11 21 31 41 51
MKTLLELLLV LLELGEAQS LHRVPLRRHP SLKKKLARS QLSEPKXSHN LDMIQFTESC 60
SMDQSAKEPL INYLMMEYFG TISIGSPPQN FTVIFDPTGSS NLWVPSVYCT SPACKTHSRF 120
QPSQSSTYSQ PQGSFSIQYQ TGLSLGIIGA QVSVESLTV VQQPGESVT EPQTFVDAE 180
FDGILGLGYP SLAVGGVTPV FDNMAQNLV DLPMSVYMS SNPEGGAGSE LIFGGYDHS 240
FSGSLNMWVPV TQAYWQIAL DNIQVGGTVM FCSEGGQAIV DTGTSITLGP SDRIKQLQNA 300
IGAAPVDGBY AVECANLNMV PDVTPTINGV PYTLSPATY LDFVDGMQF CSSGPGGLDI 360
HPPAGPLWIL GDVFIRQPSV VFDGNNRVG LAPAVP

Seq ID NO: 468 DNA sequence
Nucleic Acid Accession #: NM_018058.1
Coding sequence: 319..1575

1 11 21 31 41 51
TACGGCTGTC GGGACGGGCA GGGGAACGCC ATCGGGGTCA CAGCCTGCGA CATCGACGGG 60
GACGGCGGGG AGGAGATCTA CTTCCTCAAC ACCAATAATG CCTTCTCGGG GGTGGCCACG 120
TACACCGACA AGTGTGTTCAA GTTCGCAAT AACCGTGGG AAGACATCCT GAGCGATGAG 180
GTCAAGCTGG CCGGTGGTGT GGCACGCTC TTTGCGGAC GCTCTGTGGC CTGTGTGGAC 240
AGAAAGGGCT CTGGAGGCTA CTCTATCTAC ATTGCCAAT ACCTCTACGG TAATGTGGGC 300
CCTGATGCC TCATTGAAAT GGACCTGAG GCCAGTGACC TCTCCGGGG CATCTGGCG 360
CTCAGAGATG TGGCTGCTGA GGCTGGGGTC AGCAATATA CAGGGGGCGC AGGCGTCAGC 420
GTGGGCCCA TCTCAGCAG CAGTGCTCG GATATCTTCT GCGACAATGA GAATGGGCCT 480
AACTTCTTTT TCCACAACCG GGGCGATGGC ACCTTTGTGG ACCTGCGGC CAGTGCTGGT 540
GTGGAGGACC CCCACGAGCA TGGGCGAGGT GTCGCCCTGG CTGACTTCAA CCGTATGGC 600
AAAGTGGACA TGTCTATGG CAACTGGAAT GGGCCCCACC GCCTCTATCT GCAAAATGAGC 660
ACCAATGGGA AGGTCCGCTT CCGGGACATC GCCTCACCA AGTTCTCCAT GCCTCCCTCT 720
GTCCGACGG TCAATACCGC CGACTTTGAC AATGACAGG AGCTGGAGAT CTCTTCAAC 780
AACATTGCTT ACOGCACTC CTCAGCCAAC CGCTCTTCC GCGTATCCG TAGAGAGCAC 840
GGAGACCCCT TCATCGAGGA GCTCAATCCC GGGAGCGCTT TGGAGCCTGA GGGCGGGGCG 900
ACAGGGGGTG TGGTGACGGA CTTCAGCGGA GACGGGATGC TGGACCTCAT CTGTGCCAT 960
GGAGATGCTA TGGCTCAGCC GCTGTCCGTC TCCGGGGGCA ATCAGGGCTT CAACAACAA 1020
TGGCTCGAG TGGTGCCACG CACCGGGTGT GGGGCTTTG CAGGGGGAGC TAAGTCTGTG 1080
CTCTACACCA AGAAGAGTGG GGGCCACCTG AGGATCATCG ACGGGGGCTC AGGCTACCTG 1140
TGTGAGATGG AGCCCGTGGC ACACCTTGGC CTGGGGAAGG ATGAAGCCAG CAGTGTGGAG 1200
GTACGCTGTC CAGATGGCAA GATGGTGAAG CGGAACGTGG CCAGCGGGGA GATGAACCTA 1260

WO 02/086443

PCT/US02/12476

5
10
15

```
GTGCTGGAGA TCCTCTACCC CCGGATGAG GACACACTTC AGGACCCAGC CCCACTGGAG 1320
ACACCAATGA ATGCATCCAG TTCCCATTCG TGTGCCCTCG AGACAAGCCC GTATGTGTCA 1380
ACACCTATGG AAGCTACAGG TGCCGGACCA ACAAGAAGTG CAGTGGGGGC TACGAGCCCA 1440
ACGAGGATGG CACAGCCTGC GTGGGGACTC TCGGCCAGTC ACCGGGCCCC CGCCCCACCA 1500
CCCCACCCGC TGCTGTCTGC ACTGCCGCTG CTGCTGCCCG TGCTGGAGCT GCCACTGCTG 1560
CACCGGTCTT GTGCTGCTGC CTAGACAGTA GGGATGTAAA GGCTGGGAG GTTAAAGBAG AGCTGCGAGC 1620
CCAGCTGCTG AGCAGGGGTG GGACATGAAC CAGCGGATGG AGTCCAGCAG GGGAGTGGGA 1680
AAGTGGGCTT GTGCTGCTGC CTAGACAGTA GGGATGTAAA GGCTGGGAG GTTAAAGBAG AGCTGCGAGC 1740
CCCAAGCCCA TCCATGCACA TTACTTAGCT AACAAATAGG GAGACTCGTA AGGCCAGGCC 1800
CTGTGCTGGG CACATAGCTG TGATCACAGC AGACAGGGTC GCTGCCCTGA TGGCGCTTAC 1860
ATTCCAGTGG GTCTAATGAC CATATCTTAG GACACAGATG TGCCGAGGGA GGTGGTGTCA 1920
CTGCACAGGA AGTATGAGGA CTTTAGTGTG CTGAGTTCAA ATCCTGATTC AGGAACTCAC 1980
AAAGCTATGT GACCTTACAC CAGTCACTTA ACTTGTAGC CATCCATTAT CGCATCTGCA 2040
AAATGGGGAT TAAGAATAGA ATCTTGGGGT TAGTGTGGAG ATTAGATTAA ATGTATGTAA 2100
GACACTTGGC ACAAAACCTG GCACATAGTA AAGGCTCAAT AAAACAAGT GCCTCTCACT 2160
GGGCTTTGTC AACACGTG
```

Seq ID NO: 469 Protein sequence
Protein Accession #: NP_060528.1

20
25
30

```
1 11 21 31 41 51
MDPEASDLR GILALRDVAA EAGVSKYTG RVSVGPILS SSASDIFCDN ENGNFLPHN 60
RGDGTFFDAA ASAGVDDPHQ HGRGVALADF NRDGKVDIVY GNWNGPHRLY LQMSHKGVR 120
FRDIASPKFS MPSPVTVIT ADFDNDQLE IFFNNIAYRS SSANRLRVI RREHGDPLIE 180
ELNPGDALEP EGRGTGGVVT DFDGDMLDL ILSHGESMAQ PLSVFRNGQ FNNWLRVVP 240
RTRVGAFARG AKVLYTKKS GAHLRIIDGG SYGLCEMEFV AHFGLGKDEA SSVETWPDG 300
KMVSRNVASG EMNSVLEILY PRDETLQDP APLETPMNAS SSHSCAETS PYVSTPMEAT 360
GAGPTRSAGV ATSPTRMAQP ANGLSASHRA PAPPPLPLLL PLPLLLPLLE LPLLRSS
```

Seq ID NO: 470 DNA sequence
Nucleic Acid Accession #: AJ279016
Coding sequence: 1..1962

35
40
45
50
55
60
65
70
75
80

```
1 11 21 31 41 51
ATGTCACAGGA TGTACCGTT CTGCTGCTG CTCTGGTTTC TGCCCATCAC TGAGGGGTCC 60
CAGCGGGCTG AACCATGTT CACTGCAGTC ACCAACTCAG TTCTGCCTCC TGACTATGAC 120
AGTAATCCCA CCCAGCTCAA CTATGGTGTG GCAGTTACTG ATGTGGACCA TGATGGGGAC 180
TTTGAGATCG TGTGGCGGG GTACAATGGA CCCAACCTGG TTCTGAAGTA TGACCGGGCC 240
CAGAAGCGGC TGGTGAACAT CGCGGTGAT GAGCGCAGCT CACCCTACTA CGCGCTGCGG 300
GACCGGCAGG GGAACGCCAT CGGGGTGACA GCCTGCGACA TCGACGGGGA CGCGCGGAG 360
GAGATCTACT TCTCAACAC CAATAATGCC TTCTCGGGGG TGGCCAAGTA CACCGACAAG 420
TTGTTCAAGT TCCCAATAAA COGGTGGGAA GACATCCTGA GCGATGAGGT CAACGTGGCC 480
CGTGGTGTGG CCAGCCTCTT TGCCGGAGCG TCTGTGGCCT GTGTGGACAG AAAGGGCTCT 540
GGAGCTACT CTATCTACAT TGCCCAATTAC GCCTACGCTA ATGTGGGCCC TGATGCCCTC 600
ATTGAATGG ACCCTGAGGC CAGTGACCTC TCCCGGGGCA TTCTGGGCTC CAGAGATGTG 660
GCTGCTGAGG CTGGGTCAG CAAATATACA GGGGCGGAG GCGTCAGCGT GGGCCCCATC 720
CTCAGCAGCA GTGCTCGGA TATCTTCTGC GACAATGAGA ATGGGCTTAA CTCTCTTTC 780
CACAAACCGG GCGATGGCAC CTTTGTGGAC GCTGCGGCA GTGCTGGTGT GGACGACCCC 840
CACCAAGCAT GCGGAGGTGT CGCCCTGGCT GACTTCAACC GTGATGGCAA AGTGGACATC 900
GTCTATGGCA ACTGGAATGG CCCCCACCGC CTCTATCTGC AAATGAGCAC CCATGGGAAG 960
GTCCGCTTCC GGGACATGCG CTCACCCAAG TTCTCCATGC CCTCCCTGT CGCACGGTCT 1020
ATCACCGCG ACTTTGACAA TGACCAGGAG CTGAGATCT TCTTCAACAA CATTGCTTAC 1080
CGCAGCTCCT CAGCAACCG CCTCTTCCGC GTCATCCGTA GAGAGCACCG AGACCCCTCT 1140
ATCGAGGAGT TCAATCCCG CGAGCGCTTG GAGCGCTGAG GCGGGGGCAC AGGGGGTGTG 1200
GTGACCGACT TCGACGGAGA CGGGATGCTG GACCTCATCT TGTCCCATGG AGAGTCCATG 1260
GCTCAGCCGC TGTCCGCTCT CGGGGCAAT CAGGGCTTCA ACAACAAGT GCTGCGAGTG 1320
GTGCCACGCA CCGGTTTGG GGCCTTTGCC AGGGGAGCTA AGGTGCTGCT CTACACCAAG 1380
AAGAGTGGG CCCACCTGAG GATCATCGAC GGGGGCTCAG GCTACCTGTG TGAGATGGAG 1440
CCCGTGGCAC ACTTTGCCCT GGGGAAGGAT GAAGCCAGCA GTGTGGAGGT GAGCTGGCCA 1500
GATGGCAAGA TGGTGAAGCG GAACGTGGCC AGCGGGGAGA TGAATCTAGT GCTGGAGATC 1560
CTCTACCCCC GGGATGAGGA CACACTTCAG GACCCAGCCC CACTGGAGTG TGGCCAAGGA 1620
TTCTCCAGC AGGAAAATGG CCATTGCAAT GACACCAATG AATGCATCCA GTTCCCATTC 1680
GTGTCCAGC GAGACAAGCC CGTATGTGTC AACACCTATG GAAGCTACAG GTGCCGGACC 1740
AACAAAGAGT CGAGTCGGGG CTAOGAGCCC AACGAGGATG GCACAGCTG CGTGGGGACT 1800
CTCGGCCAGT CACCGGGCCC CGCGCCACC ACCCCACCG CTGCTGTGTC CACTGCCGCT 1860
GCTGCTGCCG GTGCTGGAGC TGCCACTGCT GCACCGTCC TCGTAGATGG AGATCTCAAT 1920
CTGGGGTCGG TGGTTAAGGA GAGCTGCGAG CCCAGCTGCT GAGCAGGGGT GGGACATGAA 1980
CCAGCGGATG GAGTCCAGCA GGGGAGTGGG AAAGTGGGCT TGTGCTGCTG CTAAGACAGT 2040
AGGGATGTAA AGGCCTGGGA GCTAGACCTC CCCCAGCCC ATCCATGAC ATTACTTAGC 2100
TAACAATTAG GGAGACTCGT AAGGCCAGGC CTTGTGCTGG GCACATAGCT GTGATCACAG 2160
CAGACAGGGT CGCTGCCCTG ATGGCGCTTA CATTCCAGTG GGTCTAATGA CCATATCTTA 2220
GGACACAGAT GTGCCAGGG AGGTGGTGTG ACTGCACAGG AAGTATGAGG ACTTTAGTGT 2280
CCTGAGTTC AATCTGATT CAGGAACCTA CAAAGCTATG TGACCTTACA CCAGTCACTT 2340
AATCTGTAG CCATCCATTA TCGCATCTCG AAAATGGGGA TTAAGAATAG AATCTTGGGG 2400
TTAGTGTGGA GATTAGATTA AATGTATGTA AGACACTTGG CACAAAACCT GGCACATAGT 2460
AAAGGCTCAA TAAACAAG TGCTCTCAC TGGCTTTGT CAACAG
```

Seq ID NO: 471 Protein sequence
Protein Accession #: CAC08451

85

```
1 11 21 31 41 51
MSRMLPFLLL LWFLPITEGS QRAEPMFTAV TNSVLPDPYD SNPTQLNYGV AVTDVDHGD 60
FEIVVAGYNG PNLVLKYDRA QKRLVNIADV ERSSPYALR DRGNAIGVT ACDIDGGRE 120
EIYFLNTNNA FSGVATYTDK LPKFRNRWE DILSDEVNVA RGVASLFAGR SVACVDRKGS 180
```

GRYSIIYIANY AYGNVGPDAL IEMDPEASDL SRGILALRDV AAEAGVSKYT GGRGVSVGPI 240
 LSSASDIFC DNENGFNLF HNRGDGTFVD AAASAGVDDP HQHGRGVALA DFNDRGKVDI 300
 VYGNWNGPHR LVLQMSHGHK VRFRDIAAPK FSPSPVVRTV ITADFDNDQ8 LEIFFNIAIY 360
 RSSANRLFR VIRREHGDPL IEELNPGDAL EPEGRGTGGV VTFDFDGMML DLILSHGESM 420
 AQLPSVFRGN QGFNNNNLWRV VPRTRFGAPA RGAKVVLYTK KSAHLRIID GSGYLCEME 480
 PVAHFLGLKD EASSVEVTPW DGKMSVRNVA SGMNSVLEI LYPRDEDTLQ DPAFLECCQG 540
 FSQOENGHCN DTNFCIOFPP VCPDRKPCV NTYGSYRCRT NKKCSRGYEP NEDGTACVGT 600
 LGQSPGPRPT TPTAAATAA AAAAAGAATA APVLVDGDLN LGSVVKESCE PSC

Seq ID NO: 472 DNA sequence
 Nucleic Acid Accession #: FGENESH
 Coding sequence: 1..4794

15 1 11 21 31 41 51
 ATGGCGTGTC CGGGAGGACT CCCAGCCCGT TGCTCTGGTT GGATGGGACT GGGTGGGCCC 60
 AGCGGCTCCT CCCAGCATC CCTCCCCAT. TCCTCCTCCA GGTACAAATG ACCCAACCTG 120
 GTTCTGAAGT. ATGACCGGGC CCAGAAGCGG CTGGTGAACA TCGCGGTGCA TGAGCGCAGC 180
 TCACCCCTACT. ACGCGCTGCG GGACCGGCAG GGAACGCCCA TCGGGGTGAC AGCCTGCGAC 240
 ATCGACGGGG ACGGCGGGGA GGAGATCTAC TTCTCAACA CCAATAATGC CTCTCGGGC 300
 CACAGCAGCT CAGCGCAGGT COCTTCTGGG CTCCACAGAA ACAGGCTGTG GCTGAAGCCT 360
 CCACCTACAA CCCCTGCAGG CCTCTGGGT CTGCTCCAC TCAGCGGAAG GGACTTTTCC 420
 TCCTCCCTGG GTACGGCTTC TCCGACAGC AGGCAGGGAG AGAGGGTGGC GGTTCCTGTC 480
 TGTCGGGGTG GACTGAGACC TACCCATGAA CCAGAACCAT TTCTTCTGAG ACCCAATCA 540
 GGGGTGGCCA CGTACACCGA CAAGTGTTC AAGTTCGCCA ATAACCGGTG GGAAGACATC 600
 CTGAGCGGATG AGGTCAAGCT. GGCCCGTGGT GTGGCCAGCT TCTTGGCCG ACCTCTGTG 660
 GCCTGTGTGG ACAGAAAGGG. CTCTGGACGC TACTCTATCT ACATTCGCAA TTACGCTTAC 720
 GGTAATGTGG GCCCTGATGC CTTATTGAA. ATGACCCCTG AGGCCAGTGA CCTCTCCCGG 780
 GGCATTCTGG. CGCTCAGAGA TGTGGCTGCT GAGGCTGGGG TCAGCAAATA TACAGAGGC 840
 TTCTCCACAC CTGCTCTCC. AAGCATTGGT GAGATATCTG GCAGAACCGA GGAGCGGGA 900
 GGAGGAGACC CAGAGGAGGC AGATGAGGAG. CACAGTGGGG ATGGAAGCAC CAGCCACTG 960
 TGCGGGCTGG GCTGGAAGGA CGGGCAGTTC AAGGAAGAA CAGCAGCTTT. GGTGGAGGAA 1020
 CAGAGGGAGG GTGGCGCAGC TGGCGTGCCC AGAGGACGTG TTCGAACAGC TCTGCAGACT 1080
 TCCAAAGGCC ATTTGGCTGA. CAAGAACCTA TTTGGCCAC CATGTTACTA TTCTGTCTGC 1140
 GCGCCTTCTC CAGCCCAACC. TTTCCCTGCC CGCCAGGCC CCCAACCTA. CCTGTAGCC 1200
 CCCCTTGTCA CTCAGCTAAT GACACATGGA CGTCTGGCTG GAAACTAGC CCGGAGTGTG 1260
 CCCACCCCC GAGCCCCAG. AATGACCCC AATGTAAGG GCGCCATGC TGAGCCCGGC. 1320
 CTGATGGCTG AGGCTTTGGG. CGCGTGGCCA GCGCTCAGCA CCACTGTGGT. GCCAGGGGGC. 1380
 CTGAGAAAGCT. GGGAGGAAAG. CAGGCAGAA GGGCAGGCCA TGTCAGATG TGCACTCAGG 1440
 GAGCTGGGAG. GTCCCTGGAG. CCAAGCCACA CAGCACCTGC CTGCTAGAGA GCTGTATGAC 1500
 CTGGGAGAAC CTCCTATTTC ACAGGAACA. GACGGAGATC CAGGGAGGAG. AAGGGACTCG. 1560
 CCCAAGGCTC CACAGGAGTG CATCTAGTG GCCACCATGC CAGCTCTCGG GGGACTCGAG. 1620
 GGCCCCGGGA GGGTGGCCAA. GCGAGAGATT GGGAGAGAGA CTGGGGCAGT AGGAAGACCA 1680
 CTCTCCCATC CCTGTGTC. CACTTCCCC AGCTGCTTGA GGCTCTTGA AGCCGGGACA. 1740
 GTGGCGGAG. CTGGCTGCC. TGGGAATCCT GGGAACTGGG. TTCTGGACAT GGCCAGGCC 1800
 CTGGCGTGA. ACCAGATGGA. AAAAGAGGAG GGAAGATT. ATGGAGACCA TGAGCCGAGA 1860
 TTTAGGCTCA GGAAGCAGC GGAAGCAGAA. TTCCCCCAG GCTCTCTGA GGAGCCTCTG 1920
 CTGCAATTC. CTTAGGCTC. CAGAGGCAGC CCTGTCTCC AGGTGGGCTT GGGGCTTGCT. 1980
 TGTGCCATC ACTGTGGGTC GATGTCTTTT CTAGGGGGCC GAGGCGTCAG CGTGGGCCCC 2040
 ATCCCTCAGCA GCAGTGCCTC GGATATCTTC TGCGACAATG. AGAATGGGCC. TAACTTCTCT. 2100
 TTCCAACACC GGGGGGATGG CACCTTTGTG GAGCGTGGG. CCAAGTGTGA ACCTGCTTTA 2160
 GCCTTCATCG TTCACTCA. ATATCACCTC TGCAAGATT TTCTCACTC CCGTGTGCCAC 2220
 CTAGCAGAA CTGGCTCTTC CTCTCTCTGC TGCCCGTGGC ATGCACGTCT. TCTTCAGGCT. 2280
 CCACATGGC ATCATGTTT. GTCTATGAGC TTACAGGA CCGGTACAG GTTCTATTCA 2340
 TTCTTACGC. AAGGCTTGGC CTCCAGTGCC CACCGGAGGA. CACTCAGCCT CCAGGGTCT. 2400
 CAGGGGGCCC CACCTTGCT. TCTGGCAAGA GCTCCCTGTG. TCTGGGGTC TGTATCCCC 2460
 ACTGCTATT. ACATGTCTCT. GTGGTCTGCC ATCCAGAGA. GCGTATGAC CCACAGCTAT 2520
 TTGTCTCTG AAGAGTCAA. CGTGGGTGTG GACGACCCCC ACCAGCATGG GCGAGGTGTC 2580
 GCGCTGGCTG ACTTCAACCC. TGATGGCAAA. GTGGACATCG TCTATGGCAA CTGGAATGGC 2640
 CCCCACCGCC. TCTATCTGCA. AATGAGCACC CATGGGAAGG TCCGCTTCCG GGACATCGCC 2700
 TCACCAAGT TCTCCATGCC CTCCCTGTG. CGCACGGTCA TCACCGCCGA. CTTTGACAAT 2760
 GACCAGAGT TGGAGTCTT CTTCACCAAC ATTGCTTACC GCAGCTCTC AGCCAACCGC 2820
 CTCTTCCGAT GCTCCATCCT GGCTCGTGGC TCTTCATCCT TGACAGCTGG TGGGAGGAAC 2880
 GGTGAGGGAG AAGGTTAAG AATCAGAAAG GAGGGTTC CAGGGCCAGG GGTGAGGCC 2940
 AAGGTCAACA CAGGTCCCT. GATGAAGAAA CAGAAAGGAA GGAAGGACGA GGAATGGGCA 3000
 AAGAGCTGTG GGAATGACAG GCAAAGCCTG GCCAAGGAGC CGGCTCTGCT TATTGAGGG 3060
 AAGGGAGAGG GAAATGTGGC CCAAAGTGTG CCGAGAACCC AAGCGCCACA AGATACAAAG 3120
 CCACACTAC ACAGAAAGGG GCTACAGGGT CCAATCACTA CCAGGAAAG GGGCTACGGG 3180
 GTCCAATCAC TACAGGAAA. AGGGGCTACG GGGTCCAATC ACTACAGGA. AAGGGGCTA 3240
 CGGGGTCCAA TCACTACAG GAAAAGGGG. TACGGGGTCC AATCACTACC AGGAAAAGGG 3300
 GCTACGGGCT CCAATCACTA CCAGGAAAG GGGCTACAGG GTCCAATCAC TACAGGAAA 3360
 AGGGGCTACG GGTCCATC. ACTACAGGA. AAGGGGCTA CAGGGTCCAA TCACTACCA 3420
 AGAAGGGGCT TACGGCTCC. AATCACTACC AGGAAAAGGG GCTACGGGGT. CCAATCACTA 3480
 CCAGGAAAG GGGCTACAGG GTCCAATCAC TACAGGAAA AGGGGCTACG GGGTCCAATC 3540
 ACTACAGGA. AAGGGGCTA. CGGGCTCAA TCACTACAG GAAAAGGGG. TACGGGGTCC 3600
 AATCACTACC AGGAAAAGGG GCTACAGGGT CCAATCACTA CCAGGAAAG GGGCTACAGG 3660
 GTCCAATCAC TACAGGAAA. AGGGGCTACG GGTCCATC. ACTACAGGA. AAGGGGCTA 3720
 CGGGGTCCAA TCACTACAG GAAAAGGGG. TACGGGGTCC AATCACTACC AGGAAAAGGG 3780
 GCTATGGGGT CCAATCACTA CCAGGAAAG GGGCTACGGT CTCCAATCAC TACAGGAAA 3840
 AGGGGCTATG GGTCCATC. ACTACAGGA. AAGGGGCTA CCGGGTCCAA CGTATCCGT 3900
 AGAGAGCAG GAGACCCCT CATCGAGGAG CTCAATCCCG GCGACGCTT. GGAGCTGAG 3960
 GGCGGGGCA CAGGGGGTGT GGTGACCGAC TTCAGCGGAG ACGGATGCT GGACCTCATC 4020
 TTGTCCCATG GAGAGTCCAT GGCTCAGCG CTGTCCGTCT TCGGGGCAA TCAGGGCTTC 4080
 AACACAACT GGCTGCGAGT GGTGCCACGC ACCCGGTTTG GGGCTTTG. CAGGGGAGCT 4140
 AAGGTGCTGC TCTACACCA. GAAGAGTGGG GCCACCTGA GATCATCGA CCGGGGCTCA 4200
 GGCTACCTGT GTGAGATGGA GCGCGTGGCA CACTTTGGCC TGGGAAGGA TGAAGCCAG 4260
 AGTGTGAGG TGACGTGGCC AGATGGCAAG ATGGTGAGCC GGAACGTGGC CAGCGGGGAG 4320

WO 02/086443

PCT/US02/12476

ATGAACCTCAG TGCTGGAGAT CCTCTACCCC CGGGATGAGG ACACACTTCA GGACCCAGCC 4380
 CCACTGGAGT GTGGCCAAAGG ATTCTCCACG CAGGAAAATG GCCATTGTCAT GGACCCAAAT 4440
 GAATGCATCC AGTTCCCATC CGTGTGCCCT CGAGACAAGC CCGTATGTGT CAACACCTAT 4500
 GGAAGCTACA GGTGCCCGAC CAACAAGAAAG TGCAGTCGGG GCTACGAGCC CAACGAGGAT 4560
 GGCAACAGCT CCGTGGGTAC TGAGCTAGGC TCTAGGCATA CAATGACGTG GAAACCAAGG 4620
 CCAAAAAGG AGCTGCAACT TTCCCAAGGC ATCTGCACCC CCGTCTGTCT CTTTTCCTG 4680
 CCGGTTGCC GGTCTCTCT CAAAAGAGCT CAGTCCAGG CTGCTCCAG CACCTTCTC 4740
 CAGAAAGCTC CAGGTATTCC AGAAGCCCAA GTGTATGAAC AAGATCAGGA ATAA

Seq ID NO: 473 Protein sequence
 Protein Accession #: FGENESH predicted

1 11 21 31 41 51
 MACPGGLPAR CSGWMGLGGP SGSSPASPPH SSSRYNGPNL VLKYDRAQKR LVNIAVDERS 60
 SPYYALDRDQ GNAIGVTACD IDGDGREEIY FLNTNNAFSG HSSSAQVPSG LHRNRPVLKP 120
 PPTTLAGLLG LPPLSGRDFS SSLGQASPDG RQGERVPVPC CRGGLRPTHE PEPFLLRPKS 180
 GVATYTRKLF KFRNNRWEDI LSDEVNVARG VASLFAGRSV ACVDRKGSOR YSIYIANYAY 240
 GNVGPDALIE MDPEASDLNR GILALRDVAA EAGVSKYTEG FSHTASPSIG EISGRTEERE 300
 GGDPEEADDEE HSGDGSTSQL CRLGKWDGQF KEEAALVEE QREAGAAGVP RGRVRTALQT 360
 SKSHLADKNL FGPPCYYSVC APSPAHPFPA RQAPQHYVPA PLVTQLMTHG RLACKLARSV 420
 PHPRAPGMDP KCKGRHAEFG LMAEALGAWP ALSTTVVPGG LRSNEESRQK GQAMSRCALR 480
 ELGGPWSQAT QHLPAELEYD LGEPPILQRT DGDPPGRRRDS PKVTQECHLV ATMPALGGLE 540
 GPGRVAKREI GRETGAVGRP LSHPLVPNFP SCLRPLEAGT VPGAALPGNF GNVRLDMAKA 600
 LAWNQMEKEE KGIHGDHEPR FRLRKAREAE FPPGSSEEP LQPPSGLRGS PVLQVGLGLA 660
 SATHCGSMNF LGGRGVSVGP ILSSASDIF CDNENGPNFL FHNRGDGTFFV DAAASAERLL 720
 APIVELKYHL CRDPFHSLSCH LAETGPSSSC CPWHARLLQA PHCHHGLSMS FTRTGSRFYS 780
 FLTQGLASSA HRRTSLQGS QGAPPCLLAR APCVLGSLIP TAYYIVLWSA IPESLMTHSY 840
 LSSERVNVGV DDPHQHGGRV ALADFNDRCK VDIVYGNWNG PHRLYLQMT HGKVRFRDIA 900
 SPKPSMPSPV RTVITADFDN DQELEIFFNN IAYRSSANR LFRCSILARG SSSLTAGGRN 960
 GQGEGLRIRR GGPFGPGQA KVTGGLMKK QKGRKDEDWA RGCGNAGQSL AKEPASAIAG 1020
 KGRGNVAGSV PRTPAQDQTK PHYHKKLGQG PITTRKRGYV VQSLPGKQAT GSNHYQEKGL 1080
 RGPITTRKRG YGVQSLPGKG ATGSNEYQEK GLQGPITTRK RGYGLQSLPG KGATGSNHYH 1140
 RKGLRAPITT RKRGYGVQSL PGKATGSNH YQEKGLRGP I TTRKRGYGLQ SLPGKATGGS 1200
 NHYQEKLGQG PITTRKRGYR VQSLPGKQAT GSNHYQEKGL RGPITTRKRG YGLQSLPGKE 1260
 AMGSNHYQEK GLRAPITTRK RGYGVQSLPG KGATGSNVIR REHGDPLIEE LNPGDALPE 1320
 GRGTGGVVD FGDGMLDLI LSHGESMAQP LSVFRGNQGF NNWLVVVR TRFGAFARGA 1380
 KVVLYTRKSG AHLRIIDGGS GYLCEMEPVA HFGLGKDEAS SVEVTWPDGK MVRNVSAGE 1440
 MNSVLEILYP RDEDTLQDPA PLECGQGSFQ QENGHOMDTP ECIQFFVFCP RDKPVCNTY 1500
 GSYRCRTNKK CSRGYPNED GTACVGTGL SRHTMTWKPR PKKELQLSQG ICTPVWSPFL 1560
 PGCRLLLKRA QLQAPSTLL QKAPGPEAQ VYEQDQE

Seq ID NO: 474 DNA sequence
 Nucleic Acid Accession #: NM_003661.1
 Coding sequence: 1..1152

1 11 21 31 41 51
 ATGAGTGCAC TTTTCTCTGG TGTGGGAGTG AGGGCAGAGG AAGCTGGAGC GAGGGTGCAA 60
 CAAAACGTTT CAGATGGGAC AGATACTGGA GATCCTCAAA GTAAGCCCT CGTGACTGG 120
 CCGTCTGGCA CCAATGGACCC AGAGAGCAGT ATCTTTATTG AGGATGCCAT TAAGTATTTC 180
 AAGGAAAAAG TGAGCACACA GAATCTGCTA CTCCTGCTCA CTGATAATGA GGCTGGAAC 240
 GGATTCTGGG CTGCTGCTGA ACTGCCAGG AATGAGGCAG ATGAGCTCCG TAAAGCTCTG 300
 GACAACTTGA CAAGACAAAT GATCATGAAA GACAAAACT GGCACGATAA AGGCCAGCAG 360
 TACAGAACT GGTITCTGAA AGAGTTTCTT CGGTTGAAAA GTGAGCTTGA GGATAACATA 420
 AGAAGGCTCC GTGCCCTTGC AGATGGGGTT CAGAAGTCC ACAAAGGCAC CACCATCGCC 480
 AATGTGGTGT CTGGCTCTCT CAGCATTTCC TCTGGCATCC TGACCTCTGT CGGCATGGGT 540
 CTGGCAGCTT TCACAGAGGG AGGCAGCCTT GTACTCTTGG AACCTGGGAT GGAGTTGGGA 600
 ATCACAGCCG CTTTGACCGG GATTACCAGC AGTACCATGG ACTACGAAA GAAGTGGTGG 660
 ACACAAGCCC AAGCCACAGA CTTGGTCATC AAAAGCCTTG ACAATTGAA GGAGGTGAGG 720
 GAGTTTTTGG GTAGAACAAT ATCCAACTTT CTTTCTTAG CTGGCAATAC TTACCAACTC 780
 ACACGAGGCA TTGGGAAGGA CATCCGTGCC CTCAGACGAG CCAGAGCCAA TCTTCAGTCA 840
 GTACCGCATG CCTCAGCCTC AOGCCCCCGG GTCAGTGAGC CAATCTCAGC TGAAGCCGT 900
 GAACAGGTGG AGAGGTTTAA TGAACCCAGC ATCCTGGAAA TGAGCAGAGG AGTCAAGCTC 960
 ACGGATGTGG CCCCTGTAG CTTCTTCTT GTGCTGGATG TAGTCTACCT CGGTACGAA 1020
 TCAAAGCACT TACATGAGGG GGCAAGTCA GAGACAGCTG AGGAGCTGAA GAAGTGGCT 1080
 CAGGAGCTGG AGGAGAAGCT AAACATTCTC AACAATAATT ATAAGATTCT GCAGGGGAGC 1140
 CAAGAAGTGT GA

Seq ID NO: 475 Protein sequence
 Protein Accession #: NP_003652.1

1 11 21 31 41 51
 MSALFLGVV RABEAGARVQ QNVPSGTDG DPQSKPLGDW AAGTMDPES IFIEDAIKYF 60
 KEKVSTQNL LLLTDNEAWN GPVAAELPR NEADELRKAL DNLARQMIMK DKNWHDKQO 120
 YRNWFLKEP RLKSELEDNI RRLRALADGV QKVHKGTTIA NVVSGSLSS SGILTIVMG 180
 LAPFTEGSSL VLEPEGMELG ITAALTGITS STMDYGKKWW TQQAQDLVI KSLDKLKEVR 240
 EFLGENISNP LSLAGNTYQL TRGICKDIRA LRRARANLQS VPHASASRPR VTEPISAESG 300
 EQVERVNEPS ILEMSRGVL TDVAPVSFFL VLDVVYLVYE SKHLHEGAKS ETAEELKKVA 360
 QELEEKLNL NNNYKILQAD QEL

Seq ID NO: 476 DNA sequence
 Nucleic Acid Accession #: NM_014452.1
 Coding sequence: 1..1968

1 11 21 31 41 51

5
10
15
20
25
30
35

ATGGGGACCT CTCCGAGCAG CAGCACCGCC CTGCGCTCCT GCAGCCGCAT CGCCCGCCGA 60
GCCACAGCCA CACATGATCG GGGCTCCCTT CTCTGCTTTG GATTCCCTAG CACCACCACA 120
GCTCAGCCAG AACAGAAGGC CTGGAATCTC ATTGGGCATC ACGCCCATGT TGACCGTGCC 180
ACCGGCCAGG TGCTAACTGT TGACAAAGTG CCAGCAGGAA CCTATGTCTC TGAGCATTTG 240
ACCAACACAA GCCTGCGCGT CTGCAGCAGT TGCCCTGTGG GGACCTTTAC CAGGCATGAG 300
AATGGCATAG AGAAATGCCA TGACTGTAGT CAGCCATGCC CATGSCCAAT GATTGAGAAA 360
TTACTTGTGT CTGCTTGTAC TGACCGAGAA TGCACTTGCC CACTTGGCAT GTTCCAGTCT 420
AACGTACTCT GTGCCGCCCA TACGGTGTGT CCTGTGGGTT GGGGTGTGCG GAAGAAAGGG 480
ACAGAGACTG AGGATGTGCG GTGTAAGCAG TGTGCTCGGG GTACCTTCTC AGATGTGCCT 540
TCTAGTGTGA TGAATGCAA AGCATAACA GACTGTCTGA GTCAGAACCT GGTGTGATC 600
AAGCCGGGGA CCAAGGAGAC AGACAACGTC TGTGGCACAC TCCCGTCTTT CTCCAGCTCC 660
ACCTCACCTT CCCCTGGCAC AGCCATCTTT CCACGCCCTG AGCAGATGGA AACCCATGAA 720
GTCCCTTCTT CCACTTATGT TCCCAAAGGC ATGAATCAA CAGAAATCAA CTCTTCTGCC 780
TCTGTTAGAC CAAAGGTACT GAGTAGCATC CAGGAAGGGA CAGTCCCTGA CAACACAAGC 840
TCCAGCAAGG GGAAGGAAGA CGTGAACAAG ACCCTCCCAA ACCTTCAGGT AGTCAACCAC 900
CAGCAAGGCC CCCACACAG ACACATCTCT AAGCTGTGCT CGTCCATGGA GGCACCTGGG 960
GGCGAGAAAT CCAGCACGCC CATCAAGGCC CCAAGAGGGG GACATCCTAG ACAGAACCTA 1020
CACAGCATT TTGACATCAA TGAGCATTTG CCTGTGATGA TTGTGCTTTT CCTGTGCTG 1080
GTGCTTGTGG TGATTGTGTG GTGCAGTATC CGGAAAGGCT CGAGGACTCT GAAAAAGGGG 1140
CCCGGCGAGG ATCCAGTGTG CATTGTGGAA AAGSCAGGGC TGAAGAAATC CATGACTCCA 1200
ACCCAGAACG GGGAGAAATG GATCTACTAC TGCAATGGCC ATGGTATCGA TATCCTGAAG 1260
CTTGTAGCAG CCAAGTGGG AAGCCAGTGG AAAGATATCT ATCAGTTTCT TTGCAATGCC 1320
AGTGAGAGGG AGGTGTCTGC TTTCTCCAAT GGTACACAG CCGACCAAGA GCGGGCCTAC 1380
GCAGCTCTGC AGCAGTGGAC CATCOGGGGC CCCGAGGCCA GCTCCGCCCA GCTAATTAGC 1440
GCCCTGCGCC AGCACCAGG AAACGATGTT GTGAGAAGA TTGTGGGCT GATGGAAGAC 1500
ACCAACCCAG TGAAGAACTGA CAAACTAGCT CTCCCGATGA GCCCAGGCC CTTAGCCCG 1560
AGCCCAATCC CCAGCCCAA CGCGAACTT GAGAATTCG CTCTCTGAC GGTGAGGCTT 1620
TCCCAACAGG ACAAGAACAA GGGCTTCTTC GTGGATGAGT CGGAGCCCTT TCTCCGCTGT 1680
GACTCTACAT CCAGCGGCTC CTCGCGCTG AGCAGGAACG GTTCTTTAT TACCAAGAA 1740
AAGAAGGACA CAGTGTGTGG GCAGGTACGC CTGGACCCCT GTGACTTGCA GCCTATCTTT 1800
GATGACATGC TCCACTTTCT AAATCTTGAG GAGCTCGGGT TGATTGAAGA GATTCCCCAG 1860
GCTGAGGACA AACTAGACCG GCTATTGGA ATATTGGAG TCAAGAGCCA GGAAGCCAGC 1920
CAGACCTTCC TGGACTCTGT TTATAGCCAT CTTCTGACC TGCTGTAG

Seq ID NO: 477 Protein sequence
Protein Accession #: NP_055267.1

40
45
50

1 11 21 31 41 51
MGTSPPSSSTA LASCSRIARR ATATMIAGSL LLLGLPLSTTT AQPEQKASNL IGTYRHVDRA 60
TGQVLTCOKC PAGTYVSEHC TMTSLRVCSS CPVGTFTPRHE NGIEKCHDCS QPCFWPMIEK 120
LPCAALTDRE CTCPPGMFQS NATCAPHTVC PVGNGVRKKG TETEDVRCKQ CARGTFSDVP 180
SSVMKCKAYT DCLSNLVVVI KPGTKETDNV CGTLPSFSSS TSPSPGTAFI PRPEHMETHE 240
VPSSTYVEKG MNTSESNSA SVRPKVLSSI QEGTVPDNTS SARGKEDVKN TLNPLQVNVH 300
QQPHHRHIL KLLPSMEATG GEKSSSTPIKG PKRGHPRQNL HKHFDINEHL PWMIVLFLLL 360
VLVVIVVCSI RKSSRTLKKG PRQDPSAIVE KAGLKKSMTP TQNRKWIYY CNHGIDILK 420
LVAAQVGSQW KDLYQFLCNA SEREVAAFSN GYTADHERAY AALQHWITRG PEASLAQLIS 480
ALRQHRNDV VEKIRGLMED TTQLETDKLA LPMSPSPSP SPSPSPNAKL ENSALLTVEP 540
SPQDKNGKFV VDESEPLLRD DSTSSGSSAL SRNGSPITKE KKDVTILQVR LDPCDLQPIF 600
DDMLHFLNPE ELRVIEEIPQ AEDKLDRLFE IIGVKQSEAS QTLDSVYSH LPDLL

Seq ID NO: 478 DNA sequence
Nucleic Acid Accession #: XM_044533
Coding sequence: 238..2751

55
60
65
70
75
80
85

1 11 21 31 41 51
GCTCTGCCCA AGCCGAGGCT GCGGGGCGGG CGCCGGCGGG AGGACTGCGG TGCCCGCGGG 60
AGGGGCTGAG TTTGCCAGGG CCACTTGAC CTTGTTTCCC ACCTCCCGCC CCCAGGTGCC 120
GGAGGCGGGG GCCCGCGGGG CGACTCGGGG GCGGACCGGG GGGCGGAGCT GCCGCGCGTG 180
AGTCCGCGCG AGCCACCTGA GCCGAGCCGG CGGGACACCG TCGCTCCTGC TCTCCGAATG 240
CTGCGCACCG CGATGGGCCT GAGGAGCTGG CTGCGCGCCC CATGGGCGGC GCTGCCGCCT 300
CGGCCACCGC TGCTGCTGCT CCTGCTGCTG CTGCTCCTGC TGCAGCCGCC GCCTCCGACC 360
TGGGCGCTCA GCCCCGGAT CAGCCTGCTT CTGGGCTCTG AAGAGCGGCC ATTCTCAGA 420
TTGGAAGCTG AACACATCTC CAACTACACA GCCCTTCTGC TGAGCAGGGA TGGCAGGACC 480
CTGTACGTGG GTGCTCGAGA GGCCTCTTT GCACTCAGTA GCAACCTCAG CTTCTGCCA 540
GGCGGGGAGT ACCAGGAGCT GCTTTGGGGT GCAGACGCG AGAAGAAACA GCACTGCAGC 600
TTCAAGGGCA AGGACCCACA GCGCGACTGT CAAACTACA TCAAGATCCT CTGCGCGCTC 660
AGCGGCACTG ACCTGTTTAC CTGTGGCACA GCAGCCTTCA GCCCATGTG TACCTACATC 720
AACATGGAGA ACTTCAACCT GGCAAGGAG GAGAAGGGGA ATGTCTCTCT GGAAGATGGC 780
AAGGGCGGTT GTCCCTTCGA CCCGAATTC AAGTCCACTG CCCTGGTGGT TGATGGCGAG 840
CTCTACACTG GAACAGTCAG CAGCTTCCAA GGGAAATGACC CGGCCATCTC GCGGAGCCAA 900
AGCCTTCGCC CACCAAGAGC CGAGAGCTCC CTCACCTGCG TGCAAGACCC AGCTTTTGTG 960
GCCTCAGCCT ACATTCTGA GAGCCTGGGC AGCTTGCAAG GCGATGATGA CAAGATCTAC 1020
TTTTTCTTCA GCGAGACTGG CCAGGAATTT GAGTTCTTGG AGAACACCAT TGTGTCCCGC 1080
ATTGCCCCGA TCTGCAAGGG CGATGAGGGT GGAGAGCGGG TGCTACAGCA CGCCTGGACC 1140
TCTTCTCTCA AGGCCAGCT GCTGTGCTCA CGGCCGAGC ATGGCTTCCC CTTCAACGTG 1200
CTGCAAGATG TCTTACGCT GAGCCCCAGC CCCAGGACT GGGGTGACAC CTTTCTCTAT 1260
GGGGTCTTCA CTTCCAGATG GCACAGGGA ACTACAGAAG GCTCTGCCGT CTGTGCTTCT 1320
ACAAATGAAG ATGTGCAGAG AGTCTTCAGC GGCCTCTACA AGGAGGTGAA CCGTGAGACA 1380
CAGCAGTGGT ACACCGTGAC CCACCCGGTG CCCACACCCC GGCTGGAGC GTGCATCACC 1440
AACAGTGCCC GGAAGAGGAA GATCAACTCA TCCCTGCAGC TCCAGACCG CGTGCTGAAC 1500
TTCTCAAGG ACCACTTCT GATGGAGCGG CAGGTCCGAA GCGCATGCT GCTGCTGACG 1560
CCCCAGGCTC GCTACAGCG CGTGGCTGTA CACCGGCTCC CTGGCTGCA CCAACCTAC 1620
GATGTCTCT TCTGGGCAC TGGTGAAGCG CGGCTCCACA AGGCAGTGAG CGTGGGCCCC 1680
CGGGTGACA CATTTAGGGA GCTGCAGATC TTCTCATCGG GACAGCCCGT GCAGAACTCG 1740

5
10
15
20
25
30
35

CTCCTGGACA CCCACAGGGG GCTGCTGTAT GGGGCTTCAC ACTCGGGCGT AGTCCAGGTG 1800
 CCCATGGCCA ACTGCAGCCT GTACAGGAGC TGTGGGGACT GCCTCCTCGC CGGGGACCCC 1860
 TACTGTGCTT GGAGCGGCTC CAGCTGCAAG CACGTACGCC TCTACAGCC TCAGCTGGCC 1920
 ACCAGGCGGT GGATCCAGGA CATCGAGGGA GCCAGCGCCA AGGACCTTTG CAGCGCGTCT 1980
 TCGGTGTGTG CCCGCTCTTT TGTACCAACA GGGGAGAAGC CATGTAGACA AGTCCAGTTC 2040
 CAGCCCAACA CAGTGAACAC TTTGGCCTGC CGCTCCTCTT CCAACTTGGC GACCCGACTC 2100
 TGGCTACGCA ACGGGGCCCC CGTCAATGCC TCGGCCTCCT GCCACGTGCT ACCCACTGGG 2160
 GACCTGTGTC TGTGGGGCAC CCAACAGCTG GGGGAGTTCC AGTGTGGTC ACTAGAGGAG 2220
 GGCTTCCAGC AGCTGGTAGC CAGCTACTGC CCAGAGGTGG TGGAGGACGG GGTGGCAGAC 2280
 CAAACAGATG AGGGTGGCAG TGTACCCGTC ATTATCAGCA CATCGCGTGT GAGTGCACCA 2340
 GCTGTGGCA AGGCCAGCTG GGTGCGAGAC AGGTCTTACT GGAAGGAGTT CCTGGTGATG 2400
 TGCACGCTCT TTTGCTGGC CGTGTGCTC CAGTCTTAT TCTTGTCTA COGGCACCGG 2460
 AACAGCATGA AAGTCTTCTT GAAGCAGGGG GAATGTGCCA CGGTGCACCC CAAGACCTGC 2520
 OCTGTGTGTC TCGCCCTGTA GACCCGCCCA CTCACCGGCC TAGGGCCCCC TAGCACCCCG 2580
 CTGCATCACC GAGGGTACCA GTCCCTGTCA GACACGCCCC CGGGGTCCCG AGTCTTCACT 2640
 GAGTCAAGA AGGGGCCACT CAGCATCCAA GACAGCTTCG TGGAGGTATC CCGAGTGTGC 2700
 CCGCGGCCCC GGGTCCGCTT TGGCTCGGAG ATCGTGACT CTGTGGTGTG AGAGCTGACT 2760
 TCCAGAGGAC GCTGCCCTGG CTTCAGGGGC TGTGAATGCT CCGAGAGGGT CCACTGGACC 2820
 TCCCTCCCGC TCTGCTCTTC GTGGAAACAG ACCGTGGTGC CCGGCCCTTG GGAGCCTTGG 2880
 GGGCAGCTGG CCTGCTGCTC TCCAGTCAAG TAGCGAAGCT CTACACCCC AGACACCCAA 2940
 ACAGCCGTGG AGCCAGAGGT CCTGGCCAAA TATGGGGGCC TGCCTAGGTT GGTGGAACAG 3000
 TGCTCCTTAT GTAAACTGAG CCTTTGTTT AAAAACAAT TCCAAATGTG AAATAGAAAT 3060
 GAGAGGGAAG AGATAGCATG GCATGCAGCA CACACGGCTG CTCAGTICA TGGCTCCCA 3120
 GGGGTGTCTG GGATGCATCC AAAGTGGTGT TCTGAGACAG AGTTGGAAC CCTCACCAAC 3180
 TGGCCTCTTC ACTTCCACA TTATCCCGCT GCCACCGGCT GCCCTGTCTC ACTGCAGATT 3240
 CAGGACCGTC SRPDDGFFN VLQDVFTLSP SPQDWRDTLP YGVFTSQWHR GTTEGSAVCV 3300
 CTGCGCTGCT CCCACCACCT CAGGGACCAAG AGGGTCAAGT TGGCACTGCG GCCCTCACA 3360
 GGTCTCTGGC TGGACCCAA CTCTGGACC TTTCCAGCCT GTATCAGGCT GTGGCCACAC 3420
 GAGAGGACAG CGGAGCTCA GGAGAGATT CTGACAAATG TACGCTTTTC CCTCAGAATT 3480
 CAGGGAAGAG ACTGTGCGCT GCCTTCTCCG GTTGTGTGCT GAGAACCCGT GTGCCCTTTC 3540
 CCACCATATC CACCTCGCTC CCATCTTTGA ACTCAAACAC GAGGAACATA CTGCACCCGT 3600
 GTCTCTTCCC CAGTCCCGAG TTCACCTTCC ATCCCTCACC TTCCTCCACT CTAAGGGATA 3660
 TCAACACTGC CCAGCACAGG GGGCCTGAAT TTATGTGTTT TTTATACATT TTTAATAAG 3720
 ATGCACTTTA TGTCATTTTT TAATAAAGTC TGAAGAATTA CTGTTT

Seq ID NO: 479 Protein sequence
 Protein Accession #: XP_044533.3

40
45
50
55

1 11 21 31 41 51
 MLRTAMGLRS WLAAPWALP PRPPLLLLLL LLLLQPPPP TWALSPRISL PLGSEERPFL 60
 RFEAEHISNY TALLLSRDR TLYVGAREAL FALSSNLSFL PGGEYQELLW GADAEEKKQC 120
 SFKGRDQPRD CQNYIKILLP LSGSHLPTCG TAAFSPMCTY INMENFTLAR DEKNVLLLED 180
 GKGRCPDPN FKSTALVVDG ELYTGTVSSF QGNDPAISRS QSLRPTKTES SLNWLDQDPAF 240
 VASAYIPESL GSLQDDDDKI YFFSETGQE PEFENTIVS RIARICKGDE GGERVLQQRW 300
 TSFLKAQILL SRPDDGFFN VLQDVFTLSP SPQDWRDTLP YGVFTSQWHR GTTEGSAVCV 360
 FTMKDQVRVF SGLYKEVNRE TQQWYVTHP VTPRPGACI TNSARERKIN SSLQLPDRVL 420
 NFLKDHFLMD GQVRSRMLL QFQARYQVVA VHRVPLHHT YDVLFLGTGD GRLEKAVSVG 480
 PRVHIIIEELQ IFSSGQPVQN LLLDTHRGLL YAASHGVVVO VPMANCSLYR SCGDCLLARD 540
 PYCAWSSSSC KHVSLYQQL ATRPWIQDIE GASAKDLCSA SSVVSPSPVP TGEKPCQEVQ 600
 PQPNTVNTLA CPLLSNLATR LMLRNGAPVN ASASCHVLPD GDLLLVTGQV LGEFQCWSLE 660
 EGFQQLVASY CPEVEDGVA DQTEGGGSPV VIISTRVSA PAGGKASWGA DRSYWKFLV 720
 MCTLFVLAVL LPVLPFLYRH RNSMKVFLKQ GECASVHPKT CPVLPPEPTR PLNGLPFPST 780
 PLDHRGYSLS SDSPPGSRVF TESEKRPLSI QDSFVEVSPV CPRPRVRLGS EIRDSVV

Seq ID NO: 480 DNA sequence
 Nucleic Acid Accession #: NM_004217.1
 Coding sequence: 58..1092

60
65
70
75
80

1 11 21 31 41 51
 GGCCGGGAGA GTAGCAGTGC CTTGGACCCC AGCTCTCCTC CCCCTTTCTC TCTAAGGATG 60
 GCCCAGAAGG AGAAGTCTTA CCCCTGGCCC TACGGCGGAC AGACGGCTCC ATCTGGCCTG 120
 AGCACCTGTC CCCAGCGAGT CCTCCGGAAA GAGCCTGTCA CCCCACTGTC ACTTGTCTTC 180
 ATGAGCGGCT CCAATGTCCA GCCACAGCT GCCCTTGGCC AGAAGGTGAT GGAGAAATAGC 240
 AGTGGGACAC CGACATCTT AACGCGGCAC TTCACAAATG ATGACTTTGA GATTGGGCGT 300
 CCTCTGGGCA AAGGCAAGTT TGGAAACGTG TACTTGGCTG GGGAGAAGAA AAGCCATTTC 360
 ATCGTGGCGC TCAAGTCTCT CTTCAAGTCC CAGATAGAGA AGGAGGGCGT GGAGCATCAG 420
 CTGCGCAGAG AGATCGAAAT CCAGGCCCCAC CTGCACCATC CCAACATCTC GCGTCTCTAG 480
 AACTATTTTT ATGACCGGAG GAGGATCTAC TTGATTCTAG AGTATGCCCC CCGCGGGGAG 540
 CTCTACAAGG AGCTGCAGAA GAGCTGCACA TTTGACGAGC AGCGAACAGC CAOGATCATG 600
 GAGGAGTTGG CAGATCTCT AATGTACTGC CATGGGAAGA AGGTGATTCA CAGAGACATA 660
 AAGCCAGAAA ATCTGCTCTT AGGGCTCAAG GAGGAGGAAG ACAATGTGTG GCACCTCGGA CTACCTGCC 720
 TCTGTGCATG CGCCCTCCCT GAGGAGGAAG GAGGAGGAAG ATCTGTGTGT CATTTGGAGT 780
 CCAGAGATGA TTGAGGGGGC CATGCACAAT GAGAAGGTGG ATCTGTGTGT CATTTGGAGT 840
 CTTTGTCTAG AGCTGCTGTT GGGGAACCCA CCTTTGAGA GTGCATCACA CAACGAGACC 900
 TATCGCGCGA TCGTCAAGGT GGACCTAAAG TTCCCGCTT CTGTGCCCCA GGGAGCCACG 960
 GACCTCATCT CCAAACTGCT CAGGCATAAC CCTCGGAAC GGCTGCCCCT GGCCAGGTC 1020
 TCAGCCCACT CTTGGGTCCG GGCCAACTCT CGGAGGGTGC TGCCTCCCTC TGCCCTTCAA 1080
 TCTGTGCGCT GATGGTCCCT GTCAATCACT CGGGTGCCTG TGTGTGTATG TCTGTGTATG 1140
 TATAGGGGAA AGAAGGGATC CTAACCTGTT CCCTTATCTG TTTTCTACCT CTTCTTTGT 1200
 TTAATAAAGG CTGAAGCTTT TTGT

Seq ID NO: 481 Protein sequence
 Protein Accession #: NP_004208

85
1 11 21 31 41 51

MAQKENSYPW PYGRQTPAPSG LSTLPQVLR KEFVTPSALV LMSRNVQPT AAPGQKVMEN 60
 SSGTPDILTR HFTIDDFEIG RPLGKGFQGN VYLAREKKSH FIVALKVLFK SQIEKEGVEH 120
 QLRREIEIOA HLHHPNLLRL YNYFYDRRRI YLILEYAPRG ELYKELQKSC TFDQRTATI 180
 MEELADALMY CHEGKVLRHD IKPENLLGLL KGELEIADFG WSWHAPSLRR KTMCGTLDYL 240
 PPENIEGRMH NEKVDLWCIG VLCYELLVGN PPFESASHNE TYRRIVKVDL KFPASVPTGA 300
 QDLISKLLRH NPSERLPLAQ VSAHPVVRAN SRRVLPPSAL QSVA

Seq ID NO: 482 DNA sequence
 Nucleic Acid Accession #: AK055663
 Coding sequence: 38..1423

1 11 21 31 41 51
 15 AGAACGGCTT CCGGCGGGAG CTGTGCAGCT CCTTATCATG GGGACAATTC ATCTCTTTG 60
 AAAACACAAA AGATCCTTTT TTGGCAAGTT GTTACGGGAA TTAGACTTGT TAGCAGCTGA 120
 CGGAAGGTCC TGGAGATAC TGCTCTTTGG TGTATAAGTA CTGGCTTCTT 180
 GCTTAGTGGG TGCAGTTCTA CTAATAGTAT AGCTTTAACT GCCTATACCT ACCTGACCAT 240
 20 TTTTGATCTT TTTAGTTTAA TGACATGTTT AATAAGTTAC TGGGTAAACAT TGAGGAAACC 300
 TAGCCCTGTC TATTGATTTG GGTTTGAAAG ATTAGAAGTC CTGGCTGTAT TGGCTCCAC 360
 AGTCTTGGCA CAGTTGGGAG CTCTCTTTAT ATTTAAAGAA AGTGCGAAGC GCTTTTGGGA 420
 ACAGCCGAG ATACACACGG GAAGATTATT AGTTGGTACT TTTGTGGCTC TTTGTTTCAA 480
 CCTGTTACAG ATGCTTTCTA TTCGGAATAA ACCTTTTGCT TATGTCTCAG AAGCTGCTAG 540
 25 TACGAGCTGG CTTCAGAGGC ATGTTGCAGA TCTTAGTCGA AGCTTGTGTG GAATTTATCC 600
 GGGACITAGC AGTATCTTCC TTCCCGAAT GAATCCATTT GTTTTATTG ATCTTGTCTG 660
 AGCATTGCTC CTTGTATTAT CATATATGCT CATTTGAAAT AATAATTATT TGGCCGTAGA 720
 CACTGCTCTC GCTATAGCTA TTGCCTTGAT GACATTTGGC ACTATGTATC CCATGAGTGT 780
 GTACAGTGGG AAAGTCTTAC TCCAGACAAC ACCACCCCAT GTTATTGGTC AGTTGGACAA 840
 30 ACTCATCAGA GAGTATCTTA CCTTAGATGG AGTTTATAGAA GTCCGAAATG AACATTTTGT 900
 GACCCTAGTC TTTGGCTCAT TGGCTGGATC AGTGCAATGA AGAATTCGAC GAGATGCCAA 960
 TGAACAAATG GTTCTGCTC ATGTGACCAA CAGGCTGTAC ACTCTAGTGT CTACTCTAAC 1020
 TGTTCAAATC TTCAAGGATG ACTGGATTAG GCCTGCCTTA TTGCTGGGCG CTGTGCGAGC 1080
 CAATGTCCTA AACTTTTCAG ATCATCAGT AATCCCAATG CCTCTTTTAA AGGCTACTGA 1140
 TGATTGGAAG CCACTTACAT CAACCTCAGC TAAACCTAGT AGTCCACCTC CAGAATTTTC 1200
 35 ATTAACTACT CCGGGGAAA ATGTGAACCC AGTTATTCTT CTAACACAC AAACAAGGCC 1260
 TTATGTTTTT GGTCTCAATC ATGGACACAC ACCTTACAGC AGCATGCTTA ATCAAGGACT 1320
 TGGAGTTCCA GGAATTGGAG CAACCTCAAG ATTGAGGACT GGTTTTACAA ATATACCAAG 1380
 TAGATATGGA ACTAATAATA GAATTGGACA ACCAAGACCA TGATAGACTC TAACCTATT 1440
 40 TTATAAGGAA TATTGACTCC TTGGCTTCCA ATTTATTAG TAATCCAAC TTGCATTGAC 1500
 TGTTTAATCA TTTACTCTAA ATGTTAGATA ATAGTAGTCT TGTTCACAT TCATGAAACC 1560
 TATGAACTA TATTTTGTGA AAATGTATT GTGACAGTGA AATCTCGTA AATGTTAAAG 1620
 GCITTTAAAT GGCTTCTCTT AGAAAATGTG TTTCTTTAAA TTTGGATTTC GGTATCTTTG 1680
 GTTTTGTAGT TGACTGCAGT GTGATGTGAC CTTACCTTTA TAAGAGCCAC TTGATGGAGT 1740
 45 AGATCTGTCA CATTACTAAG ATACGATATT TCTTTTTTT TCGAGAGCGG AGTCTTGCTC 1800
 TGCCACTGTG CCGGCCAAT ACATTATTAT TAACTTAAGG CTGTACTTTA TTAAGGCTTC 1860
 CTTAGTTTTT GTTTTGTGTT GTTTTGTGAG ATGGAGTCTC ACTCTGTGCG CCAGGCTGGA 1920
 ATGCAGTGGC ATGATCTCAG CTCACCTCAA CCTCTGCTC CTGAGTTCAA ATGATTCTCC 1980
 TGCTCAGCC TCCGAGTAG CTGGGATTAC AGGCACCTGC CACCACGCCC AGCTAATTTT 2040
 50 TGTATTTTTA GTAAAGACGG GGGATTTCAC CATGTTGGCC AGGCTGGTCT TGAACCTGCT 2100
 ACCTCATGAT CCACCACTC TAGCCTCCCA AAGTCTGGG ATTAGGTGTG AGCCACGCGA 2160
 CCTGCGCGAT ATTTTCTTTA ATGAAATTTA TAAATATGCT TCTTGAATAA TACACATTTT 2220
 GGGAAAGGGA AAAATGTCTG TTCAAAAAGT AAAGGTCTCT TTTATAGCTT TTCAAACTT 2280
 AATTGCTAAA TTTTCTTTG AGGTTCTCCT GAATTATGTC TTACAACTA AAAGCAAAA 2340
 55 TTTTAGCAG AAATTTTGA ATACATTCTA TCTAGCACAA TTTGAATTT TAATTATCAA 2400
 GATTTTTGTT AAAGTTTCTC TCCTTAAAA ATTTAGTAC ATTTGTAAAT

Seq ID NO: 483 Protein sequence
 Protein Accession #: BAB70980.1

1 11 21 31 41 51
 60 MGTIHLFRKP QRSFFGKLLR EPRVAADRR SWKILLFGVI NLICTGFLLM WCSSTNSIAL 60
 TAYTYLTIFD LFSMTCLIS YWVTLRKPS VYSFGFERLE VLAVFASTVL AOLGALFILK 120
 65 ESAERFLEQP EIHTGRLLVG TFVALCFNLF TMLSRNKPF AYVSEAASTS WLQEHVADLS 180
 RSLCGIIPGL SSIFLPRMNP FVLIDLAGAF ALCITYMLIE INNYFAVDTA SAIAIALMTF 240
 GTMYFMSVYS GKVLQITPP HVIGQLDKLI REVSTLDGVL EVRNEHFWTL GFGSLAGSVH 300
 VRIRRDANEQ MVLAVHVNRL YTLVSTLTQ IFKDDWIRPA LLSGPVAANV LNFSDDHVIP 360
 MPLLKGTDDL NPVTSTPAKP SSPPFESFN TPGKNVNFVI LLNTQTRPYG FGLNHGHTPY 420
 70 SSMLNQGLGV PGIGATQGLR TGFTNIPSRV GTNNRIGQPR P

Seq ID NO: 484 DNA sequence
 Nucleic Acid Accession #: FGENESH predicted
 Coding sequence: 1..900

1 11 21 31 41 51
 75 ATGCCGCGGC GGGAGCTGAG CGAGGCCGAG CCGCCCCCGC TCCGGGCCCC GACCCCTCCC 60
 CCGGGGCGGC GTAGCGCGCC CCCAGAGCTG GGCATCAAGT GCGTGCTGGT GGGCGACGGC 120
 80 GCCGTGGGCA AGAGCAGCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGCGCTAC 180
 CGGCCACTG CGCTGGACAC CTTCTCTGGT ACGTACGTTT AATCGCCCGT CGCGCGCGCT 240
 GGCTGCGGCG GGGCTGTGCA CCGGGGAGCT GGGCGGGGCG TCTCGGCGGG AGGGCGCAGA 300
 GGACCCCGGG GAGGAGACTG GAGCAGGCCG CGAGGTGGCG CTGGTGCGGC CCAGGACGCT 360
 CTTCTTAATC CAGGCTCTCC CGGCCCGGCC CTTGCAGTGC AAGTCTGCTG GGATGAGCT 420
 85 CCGGTGCGCA TTGAGCTCTG GGACACAGCG GGACAGGAGG ATTTTGACCG ACTTCGTTC 480
 CTTTGCTACC CGGATACCGA TGTCTTCTG GCGTCTTCA GCGGTGTGCA GCCCAGCTCC 540
 TTTCAAAACA TCACAGAGAA ATGGCTGCCC GAGATCCGCA CGCACAACCC CCAGGCGCCT 600
 GTGCTGCTGG TGGGCACCCA GGCCGACCTG AGGACGATG TCAACGTACT AATTCAGCTG 660

GACCCAGGGG GCGGGGAGGG CCCCGTGCCC CAACCCAGG CTCAGGGTCT GGCCGAGAAG 720
ATCCGAGCCT GCTGCTACCT TGAGTGCTCA GCCTTGACGC AGAAGAACTT GAAGGAAAGTA 780
TTTGACTCGG CTATTCTCAG TGCCATTGAG CACAAAGCCC GGCTGGAGAA GAAACTGAAT 840
GCCAAAGGTG TGGCAGCCCT CTCCCGCTGC CGCTGGAAGA AGTTCCTCTG CTTGCTTTGA

Seq ID NO: 485 Protein sequence
Protein Accession #: PGENESH predicted

1 11 21 31 41 51
MPPRELSEAE PPPLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY 60
RPTALDTFSG TYVQSPVRRP GCGGAVHRGA GAGVSAGRRR GPRGGDWSRP RGGAGAAQDA 120
LPNSGSPRPA PAVQVLVDGA PVRIELWDTA GOEDFDRIRS LCYPDTDVFL ACFSVVQPSS 180
FQNIETKWL P EIRTHNPQAP VLLVGTQADL RDDVNVLQL DQGGREGPVP QPQAQGLAEK 240
IRACCYLECS ALTKQNLKEV FDSAILSAT E HKARLEKRLN AKGVRTL SRC RWKPFPCFV

Seq ID NO: 486 DNA sequence
Nucleic Acid Accession #: XM_063832.2
Coding sequence: 1..711

1 11 21 31 41 51
ATGCCGCCGC GGGAGCTGAG CGAGGCCGAG CCGCCCCCGC TCCGGGCCCC GACCCCTCCC 60
CCCGCGGGGC GTAGCGCGCC CCCAGAGCTG GGCATCAAGT CGCTGCTGGT GGGGGAACGC 120
GCGGTGGGCA AGGACAGCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGCGCTAC 180
CGGCCCACTG CGCTGACAC CTTCTCTGTG CAAGTCCTGG TGGATGGAGC TCCGGTGGCG 240
ATTGAGCTCT GGGACACAGC GGGACAGGAG GATTTTGACC GACTTCGTTC CTTTGTCTAC 300
CCGGATACCG ATGCTCTCTT GCGTGTCTTC AGCGTGTGTG AGCCCACTC CTTTCAAAC 360
ATCACAGAGA AATGGCTGCC CGAGATCCGC ACGCACAAAC CCCAGGCGCC TGTGCTGCTG 420
GTGGGCAACC AGGCCAGCCT GAGGAGCAT GTCAACGTAC TAATTCAGT GGACCAAGGG 480
GCCCGGAGGG GCCCGTGGCC CCAACCCAG GCTCAGGGTC TGGCCGAGAA GATCCGAGCC 540
TGCTGCTACC TTGAGTGCTC AGCCTTGACG CAGAAGAACT TGAAGGAAGT ATTTGACTCG 600
GCTATTCTCA GTGCCATTGA GCACAAAGCC CGCTGGAAGA AGAAGCTGAA TGCCAAAGGT 660
GTGCGCACCC TCTCCCGCTG CGCTGGAAG AAGTTCTTCT GCTTCGTTTG A

Seq ID NO: 487 Protein sequence
Protein Accession #: XP_063832.1

1 11 21 31 41 51
MPPRELSEAE PPPLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY 60
RPTALDTFSV QVLVDGAPVR IELMDTAGQE DFDRRLSLCY PDTDVFLACF SVVQPSFQFN 120
ITEKWLPEIR THNPQAPVLL VGTQADLRDD VNVLIQLDQG GREGPVPQPO AQGLAEKIRA 180
CCYLECSALT QKNLKEVFDS AILSAIEHKA RLEKKLNAGK VRTL SRCRWK KFPFCFV

Seq ID NO: 488 DNA sequence
Nucleic Acid Accession #: NM_014398.1
Coding sequence: 64..1314

1 11 21 31 41 51
GGCACCAGAT CGGGGCTGCG CGGACTTCG CCGCACGCTG CAGAACCTCG CCCAGCGCCC 60
ACCATGCCCC GGCAGCTCAG CGGCTCTTCG CGCTCTTCGG CGTCCCTGGC CGTAATTTTG 120
CAGCATGGCA GTCAAAATGAG AGCAAAAGCA TTTCCAGAAA CCAGAGATTA TTCTCAACCT 180
ACTGCAGCAG CAACAGTACA GGACATAAAA AAACCTGTCC AGCAACCCAG TAAGCAAGCA 240
CCTCACCAAA CTTTAGCAGC AAGATTTCATG GATGGTCATA TCACCTTTCA AACAGCGGCC 300
ACAGTAAAAA TTCCACAAC TACCCAGCA ACTACAAAAA ACACCTGCAAC CACCAGCCCC 360
ATTACCTACA CCCTGCTCAG AACCCAGGCC ACACCCAACT ACTCACACAC AGCTCCTCCA 420
GTTACTGAAG TTACAGTCGG CCCTAGCTTA GCCCCTTATT CACTGCCACC CACCATCAAC 480
CCACCACTC ATACAGCTGG AACCACTTCA TCAACCGTCA GCCACACAAC TGGGAACACC 540
ACTCAACCCA GTAAACAGAG CACCCCTCCA GCAACTTTAT CGATAGCACT GCACAAAAGC 600
ACAACCGGTC AGAAGCCTGA TCAACCCACC CATGCCCCAG GAACCAAGGC AGCTGCCACC 660
AATACCAACC GCACAGCTGC ACCTGCCTCC ACGGTTCTCG GGCCCAACCT TGCACCTCAG 720
CCATCGTCAG TCAAGACTGG AATTATCAG GTTCTAAACG GAAGCAGACT CTGTATAAAA 780
GCAGAGATGG GGATACAGCT GATTGTTCAA GACAGGAGT CGGTTTTTTC ACCTCGGAGA 840
TACTTCAACA TCGACCCCAA CGCAACGCAA GCCTCTGGGA ACTGTGGCAC CGGAAAATCC 900
AACCTTCTGT TGAATTTTCA GGGCGGATT GTGAATCTCA CATTTACCAA GGATGAAGAA 960
TCATATTATA TCAGTGAAGT GGGAGCCTAT TTGACCGTCT CAGATCCAGA GACAGTTTAC 1020
CAAGGAATCA AACATGCGGT GGTGATGTTT CAGACAGCAG TGGGCAATTC CTTCAAGTGC 1080
GTGAGTGAAC AGAGCCTCCA GTTGTGAGCC CACCTGCAGG TGAACAACAC CGATGTCCAA 1140
CTTCAAGCCT TTGATTTTGA AGATGACCAC TTTGGAAATG TGGATGAGTG CTCGTCTGAC 1200
TACACAATTG TGCTTCCTGT GATTGGGGCC ATCGTGGTTG GTCTCTGCTT TATGGGTATG 1260
GGTGTCTATA AAATCCGCCCT AAGTGTCAA TCATCTGGAT ACCAGAGAAT CTAATTGTTG 1320
CCCGGGGGGA ATGAAAATAA TGAATTTAG AGAAGCTTTT CATCCCTTCC AGGATGGATG 1380
TTGGGAAATT CCCTCAGAGT GTGGGTCTTT CAAACAATGT AAACCAACAT CTTCTATTCA 1440
AATGAAGTGA GTCATGTGTG ATTAAAGTTC AGGCAGCACA TCAATTTCTA AATACTTTT 1500
GTTTATTTTA TGAAGATAT AGTGAGCTGT TTAATTTCTA GTTTCCTTTA GAATATTITA 1560
GCCACTCAAA GTCAACATTT GAGATATGTT GAATTAACAT AATATATGTA AAGTAGAATA 1620
AGCCTTCAAA TTATAACCA AGGGTCAATT GTAACATAA CTACTGTGTG TGCAATTGAG 1680
ATTTTATTTT ACCCTGATC TTAACAAAGC CTTTGTCTTG TTATCAATG GACTTTCACT 1740
GCTTTACTCA TCTGTGTTT ATGGTTTCAT GTAACATAA TATTCCTGGT GTAGCACTTA 1800
ACTCCTTTTC CACTTTAAAT TGTGTTTGT TTTTGGAGC GGAGTTTCC TCTTGTCAAC 1860
CAGGCTGGAG TACAGTGCCA CGATCTCGGC TTATGGCAAC CTCGCTTCC CGGGTTCAAG 1920
TGATTCTCCT GCTTCAGCTT CCGAGTAGC TGGGATTACA GGCACACACT ACCACGCTG 1980
GCTAATTTT GTATTTTAT TATAGACGGG TTTCAACATG TTGGCCAGC TGGTCTTGAA 2040
CTCTGACCT CAGGTGATCC ACCCACTCA GCCTCCCAA GTGCTGGGAT TACAGGCAT 2100
AGCCATTGCG CCCGCCCTTA AATGTTTTT TTAATCATCA AAAAGAACA CATATCTAG 2160

GTTGTCTAAG TGTTTTATG TAAACCAAC AAAAGAACA AATCAGCTTA TATTTTTTAT 2220
 CTGTGAGACT CCTGCTCCAG AATTGCTAGA CTAAGAATTA GGTGGCTACA GATGGTAGAA 2280
 CTAACAATA AGCAAGAGAC AATAATAATG GCCCTTAATT ATTAACAAAG TGCCAGAGTC 2340
 TAGGCTAAGC ACTTATCTTA TATCTCATTT CATTTCTACA ACTTATAAGT GAATGAGTAA 2400
 5 AGTGAAGACT AAGGGAACCT AATCACTTAA ATGTCACCTG GCTAACTGAT GGCAGAGCCA 2460
 GAGCTTGAAT TCATGTTGGT CTGACATCAA GGTCTTTGGT CTCTCCCTA CACCAAGTTA 2520
 CCTACAAGAA CAATGACACC ACACCTCGCC TGAAGGCTCA CACCTCATAC CAGCATACGC 2580
 TCACCTTACA GGGAAATGGG TTTATCCAGG ATCATGAGAC ATTAGGGTAG ATGAAGGAG 2640
 AGCTTTGACG ATAACAAAAT AGCCTATCCT TAATAAATCC TCCACTCTCT GGAAGGAGAC 2700
 10 TGAGGGGCTT TGTAAAACAT TAGTCAGTTG CTCATTTTTA TGGGATTGCT TAGCTGGGCT 2760
 GTAAAGATGA AGGCATCAAA TAAACTCAAA GTATTTTTAA ATTTTTTTGA TAATAGAGAA 2820
 ACTTCGTGTA CCAACTGTTC TTTCTTGAGT GTATAGCCCC ATCTTGTGGT AACTTGCTGC 2880
 TTCTGCACTT CATATCCATA TTTCTATTG TTCACCTTAT TCTGTAGAGC AGCCTGCCAA 2940
 GAATTTTTAT TCTGCTGTTT TTTTGTCTGC TAAAGAAAGG AACTAAGTCA GGATGTTAAC 3000
 15 AGAAAGTCC ACATAACCCT AGAATCTTAA GTCAAGGAAT AATTCAAGTC AGCCTAGAGA 3060
 CCATGTTGAC TTTCTCTATG TGTTCCTTAA TGACTCAGTA AGTTGGCAGG GTCTGACTT 3120
 TAGTCTTAAT AAAACATTGA ATTGTAGTAA AGGTTTTTGC AATAAAAACT TACTTTGG

Seq ID NO: 489 Protein sequence
 Protein Accession #: NP_055213.1

1 11 21 31 41 51
 MPRQLSAAAA LFASLAVILH DGSQMRKAF PETRDYSQPT AAATVQDIKK PVQPPAKQAP 60
 25 HQTLAARFMD GHITFQTAAT VKIPTTTPTAT TKNTATTSPI TYTLVTTQAT PNNSTAPPV 120
 TEVTVGPSLA PYSLPPTITP PAHTAGTSSS TVSHTTGNTT QPSNQITLPA TLSIALHKST 180
 TGQKPDQPTH APGTTAAAHN TTRTAAPAST VPGHTLAPQP SSVKTGIYQV LNSRLCIKA 240
 EMGIQLIVQD KESVFSPPRY FNIDPNATQA SCNCGRKSN LLLNFQGGPV NLTFKDEES 300
 30 YYISEVGAYL TVSDPEVTYQ GIKHAVVMPQ TAVGHSFKCV SEQSLQLSAH LQVKTDDVQL 360
 QAFDFEDDFH GNVDECSSDY TIVLPVIGAI VVGLCLMGMG VYKIRLRCS SGYQRI

Seq ID NO: 490 DNA sequence
 Nucleic Acid Accession #: NM_005409.3
 Coding sequence: 94..378

1 11 21 31 41 51
 TTCCTTTCAT GTTCAGCATT TCTACTCCTT CCAAGAAGAG CAGCAAAGCT GAAGTAGCAG 60
 40 CAACAGCACC AGCAGCAACA GCAAAAAACA AACATGAGTG TGAAGGGCAT GGCTATAGCC 120
 TTGGCTGTGA TATTGTGTGC TACAGTTGTT CAAGGCTTCC CCATGTTCAA AAGAGGACGC 180
 TGTCTTTTGA TAGGCCTCGG GGTAAAAGCA GTGAAAGTGG CAGATATTGA GAAAGCCTCC 240
 ATAATGTACC CAAGTAAACA CTGTGACAAA ATAGAAGTGA TTATTACCTT GAAAGAAAAT 300
 45 AAAGGACAAAC GATGCCATAA TCCCAATCC AAGCAAGCAA GGCTTATAAT CAAAAAAGTT 360
 GAAAGAAAGA ATTTTAAAAA ATATCAAAAC ATATGAAGTC CTGGAAAAGG GCATCTGAAA 420
 AACCTAGAAC AAGTTTAACT GTGACTACTG AAATGACAAG AATTCTACAG TAGGAAACTG 480
 AGACTTTTCT ATGTTTGTGT GACTTTCAAC TTTTGTACAG TTATGTGAAG GATGAAAGGT 540
 GGGTGAAAGG ACCAAAACAA GAAATACAGT CTCTCTGAAT GAATGACAAT CAGAATTCCA 600
 CTGCCCAAAG GAGTCCAGCA ATTAATGGA TTTCTAGGAA AAGCTACCTT AAGAAAGGCT 660
 50 GGTTACCATC GGAGTTTACA AAGTGCTTTC ACGTTCTTAC TTGTTGTATT ATACATTCAT 720
 GCATTTCTAG GCTAGAGAAC CTTCTAGATT TGATGCTTAC AACTATTCTG TTGTGACTAT 780
 GAGAACATTT CTGTCTCTAG AAGTTATCTG TCTGTATTGA TCTTTATGCT ATATTACTAT 840
 CTGTGGTTAC AGTGGAGACA TTGACATTAT TACTGGAGTC AAGCCCTTAT AAGTCAAAAG 900
 CATCTATGTG TCGTAAAGCA TTCTCAAAAC ATTTTTCAT GCAAAATACAC ACTCTTTCC 960
 55 CCAAAATATCA TGTAGACAT CAATATGTAG GAAAACATTC TTATGCATCA TTTGGTTTGT 1020
 TTTATAACGA ATTCATTAAA GTTAATTCTAT AAAATGTACT ATGAAAAAAA TTATACGCTA 1080
 TGGGATACAG GCACAGTGC ACATATTCA TAAACAAAT AGCAGCACCG GTCTTAATTT 1140
 GATGTTTTTC AACTTTTATT CATTGAGATG TTTTGAAGCA ATTAGGATAT GTGTGTTTAT 1200
 TGTACTTTTT GTTTTGATCC GTTTGTATAA ATGATAGCAA TATCTTGGAC ACATTGAAA 1260
 60 TACAAATGT TTTTGTCTAC CAAAGAAAAA TGTTGAAAAA TAAGCAAATG TATACCTAGC 1320
 AATCACTTTT ACTTTTGTGA ATTCTGTCTC TTAGAAAAAT ACATAATCTA ATCAATTTCT 1380
 TTGTCATGTC CTATATACAG TAAAAATTAG GTATACTCAA GACTAGTTTA AAGATCAAA 1440
 GTCATTTTTT TCTCTAATAA ACTACCACAA CCTTTCTTTT TTAAAAAAAA AAA

Seq ID NO: 491 Protein sequence
 Protein Accession #: NP_005400.1

1 11 21 31 41 51
 MSVKGMALAL AVILCATVVQ GFPMPKRGRC LCIGPGVKAV KVADIEKASI MYPNNNDKI 60
 EVIITLKENK GQRCLNPKSK QARLIILKVE RKNF

Seq ID NO: 492 DNA sequence
 Nucleic Acid Accession #: NM_000577.1
 Coding sequence: 41..520

1 11 21 31 41 51
 GGCACGAGGG GAAGACCTCC TGTCTATCA GGCCTCCCC ATGGCTTTAG AGACGATCTG 60
 80 CCGACCTCTT GGGAGAAAAA CCAGCAAGAT GCAAGCCTTC AGAATCTGGG ATGTTAACCA 120
 GAAGACCTTC TATCTGAGGA ACAACCAACT AGTTGCCGGA TACTTGCAAG GACCAATGT 180
 CAATTTAGAA GAAAGATAG ATGTGGTACC CATTGAGCCT CATGCTCTGT TCTTGGGAAT 240
 CCATGGAGGG AAGATGTGCT TGTCTGTGT CAAGTCTGGT GATGAGACCA GACTCCAGCT 300
 GGAGGCAGTT AACATCACTG ACCTGAGCGA GAACAGAAAG CAGGACAAGC GCTTCGCCTT 360
 CATCCGCTCA GACATGGGCC CCACCAACAG TTTTGTAGTCT GCCGCTGCC CCGTGTGGTT 420
 85 CCTCTGCACA GCGATGGAAG CTGACCAAGC CGTCAGCCTC ACCAATATGC CTGACGAAGG 480
 CGTCATGGTC ACCAAATTCT ACTTCCAGGA GGACGAGTAG TACTGCCACG GCCTGCCTGT 540
 TCCCATTTCT GCATGGCAAG GACTGCAGGG ACTGCCAGTC CCCCTGCCCC AGGGCTCCCG 600

5
10
15
20

```

GCTATGGGG CACTGAGGAC CAGCCATTGA GGGGTGGACC CTCAGAAGGC GTCACAACAA 660
CCTGGTCACT GGACTCTGAC TCCTCTTCAA CTGACCAGCC TCCATGCTGC CTCCAGAATG 720
GTCTTTCTAA TGTGTGAATC AGAGCACAGC AGCCCTGCA CAAAGCCCTT CCATGTCGCC 780
TCTGCATICA GGATCAAAAC CCGACCACCT GCCCAACCTG CTCTCCTCTT GCCACTGCCT 840
CTTCTCCCTT CATTCCACCT TCCCATGCC TGGATCCATC AGGCCACTTG ATGACCCCA 900
ACCAAGTGGC TCCACACACC TGTTTTACAA AAAAGAAAAG ACCAGTCCAT GAGGGAGGTT 960
TTTAAGGTT TGTGAAAAT GAAAATTAGG ATTTTCATGAT TTTTTTTTTT CAGTCCCCGT 1020
GAAGGAGAGC CCTTCATTGT GAGATTATGT TCTTTCGGGG AGAGGCTGAG GACTTAAAT 1080
ATTCTGTCAT TTGTGAAATG ATGGTGAAAG TAAGTGTAG CTCTTCCCTT CTTTTCTTC 1140
TTTTTTTGTG ATGTCCCAAC TTGTAAAAAT TAAAGTTAT GGTACTATGT TAGCCCCATA 1200
ATTTTTTTTT TCCTTTTAAA ACACTTCCAT AATCTGGACT CCTCTGTCCA GGCATGCTG 1260
CCCAGCCTCC AAGCTCCATC TCCACTCCAG ATTTTTTACA GCTGCTGCA GTACTTTACC 1320
TCCTATCAGA AGTTTCTCAG CTCCCAAGGC TCTGAGCAA TGTGGCTCCT GGGGTTCTT 1380
TCTTCTCTG CTGAAGGAAT AAATTGCTCC TTGACATTGT AGAGCTCTG GCACTTGGAG 1440
ACTTGTATGA AAGATGGCTG TGCCCTGCTC GTCTCCCTCC ACCAGGCTGG GAGCTCTGCA 1500
GACGAGGAAA CATGACTCGT ATATGTCTCA GGTCCCTGCA GGGCCAGCA CCTAGCCTCG 1560
CTCTTGGCAG GTACTACAGC AATGAATGCT GTATATGTTG GGTGCAAGT TCCTACTTTC 1620
CTGTGACTTC AGCTCTGTTT TACAATAAAA TCTTGAATAA GCCTAAAAAA AAAAAAAA 1680
AAAAAAA AAAA AAAA AAAA

```

Seq ID NO: 493 Protein sequence
Protein Accession #: NP_000568.1

25

```

1 11 21 31 41 51
MALETICRPS GRKSSKMQAF RIWDVNQKTF YLRNNQLVAG YLQGPVNMLE EKIDVVP IEP 60
HALFLIGHG KMCLSCVKSQ DETRIQLQLEAV NITDLSENK QDKRFAPFIR DSGPTTSFES 120
AACPGWFLCT AMEADQPVSL TNMPDEGVMV TKFYFQEDE

```

30
Seq ID NO: 494 DNA sequence
Nucleic Acid Accession #: NM_002081.1
Coding sequence: 222..1898

35
40
45
50
55
60
65
70
75
80
85

```

1 11 21 31 41 51
GGCTGCCCGA GCGAGCGTTC GGACCTCGCA CCCCGCGCGC CCCGCGCCGC CGCGCGCCGC 60
GGCTTTTGTG GTCTCGCGCT CCTCGGCGCG CGCGCGCTCT GGACCGCGAG CCGCGCGCGC 120
CGGGACCTTG GCTCTGCGCT TCGCGGCGCG GAACTGCGCA GGACCGCGCC AGGATCCGAG 180
AGAGCGCGCG GCGGGTGGCC GGGGCGCGCG CGCGCGCGCG CATGAGCTC CGGGCGCGAG 240
GCTGGTGGCT GCTATGTGGC GCGCGAGCGC TGGTGCCTG CGCGCGCGCG GACCGCGCCA 300
CGAAGAGCGG GAGCTGCGCG GAGGTCCGCC AGATCTACGG AGCCAGGGCG TTGACGCTGA 360
GCGAGCTGCC CCAGCGCGAG ATCTCGGGTG AGCACTCGCG GATCTGTCCC CAGGGCTACA 420
CCTGCTGCAC CAGCGAGATG GAGGAGAACC TGGCCAACCG CAGCATGCC GAGCTGGAGA 480
CGCGCTCCG GGACAGCAGC CGCGTCTGCG AGGCCATGCT TGCCACCCAG CTGCGCAGCT 540
TCGATGACCA CTTCACGAC CTGCTGAACG ACTCGGAGCG GACGCTGCAG GCCACCTTCC 600
CGCGCGCTTT CGGAGAGCTG TACACGCGAG ACGCGAGGGC CTTCCGGGAG CTGTACTCAG 660
AGCTGCGCCT GTACTACGCG GGTGCCAACC TGCACTGGA GGAGACGCTG GCGGAGTTCT 720
GGGCGCGCCT GCTCGAGCGC CTCTCAAGC AGCTGCACCC CAGCTGCTG CTGCGTATG 780
ACTACCTGGA CTGCTCGGGC AAGCAGGCGG AGCGCTGCG GCGCTTCCGG GAGGCGCCGA 840
GAGAGCTGGC CTTGCGGGCC ACCCGTGCTT TGGTGGCTG TCGCTCTCTT GTGCAAGGCC 900
TGGCGTGGC CAGCGAGCTG GTCCGGAAG TGGCTCAGT CCGCTTGGCG CCGGAGTGCT 960
CGAGAGCTGT CATGAAGCTG GTCTACTGTG CTCACTGCTT GGGAGTCCCC GCGCGCAGCG 1020
CCTGCCCTGA CTATTGCCGA AATGTGCTCA AGGCTGCTC TGCCAACCA GCGGACCTGG 1080
ACGCGAGTGT GAGGAACCTC CTGGACTCCA TGGTGTCTAT CACCGACAAG TTCTGGGGTA 1140
CATCGGGTGT GGAGAGTGTG ATCGGCGAGC TGCAACAGTG GCTGGCGSAG GCCATCAACG 1200
CCCTCCAGCA CAACAGGAGC ACGCTCACGG CCAAGGTCTT CCAGGGCTG GGAACCCCA 1260
AGGTCAACCC CCAGGGCCCT GGGCTGAGG AGAAGCGCG CCGGGGCAAG CTGCGCCCGC 1320
GGGAGAGGCC ACCTTCAGGC ACGCTGGAGA AGCTGGTCTC TGAAGCCAAG GCCAGCTCC 1380
GCGAGCTCCA GGACTCTGAG ATCAGCCTCC CAGGGAGACT GTGCACTGAG AAGATGGCCC 1440
TGAGCACTGC CAGTGTATG CCGTGTCTGA ACGGGATGGC CAGAGGCGCG TACCTCCCCG 1500
AGGTATGTC TGACGGCTCG GCCAACCA TCAACAACCC CGAGGTGGAG GTGGACATCA 1560
CCAAGCGGGA CATGACATC CCGCAGCAGA TCATGCAGCT GAAGATCATG ACCAACCGGC 1620
TGCGCAGCGC CTACACCGC AACGACGTGG ACTTCCAGGA CGCCAGTGAC GACGCGCAGG 1680
CGTGGGCGAG CGGTGATGGC TGTCTGGATG ACCTCTGCGG CCGGAAGGTC AGCAGGAAGA 1740
GCTCCAGCTC CCGGACGCCC TTGACCCATG CCTTCCAGG CCTGTGAGC CAGGAAGGAC 1800
AGAAGACCTC GGCTGCCAGC TGCCCCCAGC CCGCGACCTT CCTCTGCCC CTCTCTCTCT 1860
TGCTGGCCCT TACAGTAGCG AGGCCCCGGT GCGGTAACCT GCGGCAAGGC CCCAGGGACA 1920
GAGGCGAAGC ACTGACTTTG CCAAAAATAC AACACAGAG ATATTAAAT CACTCAGGCC 1980
TGGAGAGGCC TGGGGTGGGA CAGGGAGGGC CGCGGCTCT GAGCAGGGGC AGGCGCAGAG 2040
GTCCAGCCCC CAGGCTGGC CTGCGCTGCC TTTCTGCTT TTAATTTTGT ATGAGGTCTT 2100
CAGGTGAGT GGGAGCCAGT GTGCCCAAAA GCCATGTATT TCAGGAGACT CAGGGGACCC 2160
TCCGCTGCTC TAGCCCTCCC CCGAGCTCCC TGCAACGCGG CAGAAGCAGC CCTCGAGGC 2220
CTACAGAGGA GGCTCAAAG CAACCGCTG GAGCCACAG CGAGCTGTG CTTCTCTCCC 2280
CGCTCTCTCC CACTGGGACT CCCAGCAGAG CCCACAGCC AGCCTTGCC CACCCCCAG 2340
CCTCCAGAGA AGCCCCGAC GGGCTGTCTG GGTGTCTGCC ATCCAGGGTC TGCGAGAGCC 2400
TCTGAGATGA TGATGATGCT CTTCCCTCA GCGCAGGCTG CAGAGCCCGC CCCCACCTCC 2460
CTGCGCCCTT GAGGGGCCCC AGCGTCTGCA GGTGAGCGCC TGAGACAGCA CCACTGCTGA 2520
GGAGTCTGAG GACTGTCTCT CCACAGACC TGCAGTGAGG GGCCCTCCAT GCGCAGATGA 2580
GGGGCACTG ACCACCTGC GCTTCTGCTG GAGGAGGGGA AGCTGGGCC AAAGGCCAG 2640
GGAGGACGCG TGGGCTCTGC CAATGTGGGC TGCCCCCTG ACACAGGGCT CACAGGGCAG 2700
GCTTGTCTG GGTTCAGGGC TGTGGAGGA CCGGAGGGC TGAGGAGCAG CAGGACCGC 2760
CCTGCTCCCA TCCTCACCA GATCAGGAAC CAGGGCTCC CTGTTACGG TGACACAGGT 2820
CAGGGCTCAG AGTGACCTTC GGCTGTACCC TGCTCACAG GATGCTGGT GCTGTGAGA 2880
CCCCGACTG CACACGGGAA TGCTAGGTC CTTCCCGAG CCAGCCAGCT GCACTGCAGG 2940
GCACGGGGAG CTGGATAGTT AAGGGCTTTT CCAACATGC ATCCATTAC TGCACTTCC 3000
TGTCCTGTT CATGAGAGC GTTCTGCTCC TCCAGATGG CTTGGAGGC CCGCAGGGCC 3060
CACCTTGGAC CTTGGTGACC TCCTGTCACT CACTGAGGCC ATCAGGGCCC TGCCCCAGGC 3120

```

WO 02/086443

PCT/US02/12476

CTGGACGGGC CCTCCTTCCC TCCTGTGCCC CAGCTGCCAG GTGGCCCTGG GGAGGGGTGG 3180
 TGTGGTGTG GGAAGGGGTC CTGCAGGGGG AGGAGGACTT GGAGGGTCTG GGGGCAGCTG 3240
 TCCTGAACCG ACTGAACCTG AGGAGGGCGC TTAGTGCTGC TTGCTTTTC ATCACCGTCC 3300
 CGCACAGTGG ACAGAGGTCC CGGTTGCTG GTGAGTCCCC CATGGCTTGT TCTCTGGAAC 3360
 CTGACTTTAG ATGTTTGTGG ATCAGGAGCC CCAACACAG GCAAGTCCAC CCCATAATAA 3420
 CCTGCGCAGT GCCAGGGTGG GCTGGGGACT CTGGCACAGT GATGCGGGC GCCAGGACAG 3480
 CAGCACTCCC GCTGCACACA GACGGCCTAG GGGTGGCGCT CAGACCCAC CCTACGCTCA 3540
 TCTCTGGAAG GGGCAGCCCT GAGTGGTCA TGTTCAGGGC AGTGGCCAG CCTGCTGTGT 3600
 CCTTCTCCA CAAGGTCCCC CCACCGCTCA GTGTCAAGGG GTGAGGTGTG TTCTTTTGAG 3660
 TCCTTGTATG AATAAAGGC TGAACCTA AA

Seq ID NO: 495 Protein sequence
 Protein Accession #: NP_002072.1

1 11 21 31 41 51
 MELRARGWNL LCAAAALVAC ARGDPASKSR SCGEVRQIYG AKGFSLSQVP QAEISGEHLR 60
 ICPOGYTCCT SEMEENLANR SHALELTALR DSSRVQLAML ATQLRSPDDH FQHLLNDSER 120
 TLQATPFGAF GELVYQNAAR FRDLYSELRL YYRGANLHLE ETLAEFWARL LERLFKQLHP 180
 QLLLPDDYLD CLGKQABALR PFGEAPRELRL LRATRAFVA RSVQGLGVA SDVVRKVAQV 240
 PLGPECSRAR MKLVYCAHCL GVPGARPCPD YCRNVLKGL ANQADLDAEW RNLLDSMVL 300
 TDKFGTSGV ESIVGSVHTW LAEAINALQD NRDLTLARVI QGCGNPKVNP QGPGPEEKRR 360
 RGKLAPRRP PSGLTLEKLV EAKAQLRDVQ DFNISLPGLT CSEKMASTA SDDRCSWNGMA 420
 RGRYLPVVMG DGLANQINNP EVEVDITKPD MTIRQQIMQL KIMTNRLRSA YNGNDVDFQD 480
 ASDDGSGSGS

Seq ID NO: 496 DNA sequence
 Nucleic Acid Accession #: NM_001650.2
 Coding sequence: 40.1011

1 11 21 31 41 51
 GGGGCAGGCA ATGAGAGCTG CACTCTGGCT GGGGAAGGCA TGAGTGACAG ACCCAGCA 60
 AGCGGGTGGG GTAAGTGTGG ACCTTTGTGT ACCAGAGAGA ACATCATGGT GGCTTTCAA 120
 GGGGTCTGGA CTCAAGCTTT CTGGAAGCA GTACAGCGG AATTCTGGC CATGCTTATT 180
 TTGTTCTCC TCAGCCTGGG ATCCACCATC AACTGGGGT GAACAGAAA GCCTTTACCG 240
 GTCGACATGG TTCTCATCTC CTTTGTCTTT GGAATCAGCA TTGCAACCAT GGTGCGTGC 300
 TTGCGCCATA TCAGCGGTGG CCACATCAAC CCTGCAGTGA CTGTGCGCAT GGTGTGCACC 360
 AGGAAGATCA GCATCGCCAA GTCTGTCTTC TACATCGCAG CCCAGTGCCT GGGGGCCATC 420
 ATTGGAGCAG GAATCCTCTA TCTGGTCACA CCTCCAGTGG TGGTGGGAGG CTTGGGAGTC 480
 ACCATGGTTC ATGGAAATCT TACCGCTGGT CATGGTCTCC TGGTGGAGTT GATAATCACA 540
 TTCAATTTGG TGTCTACTAT CTTTGCAGCG TGTGATTCCA AACGAGTGA TGTCACTGGC 600
 TCAATAGCTT TAGCAATTGG AATTTCTGTT GCAATTGGAC ATTTATTTC AATCAATTAT 660
 ACTGTGCGCA GCATGAATCC CGCCCGATCC TTTGGACCTG CAGTTATCAT GGGAAATTGG 720
 GAAAACCAT TGGATATATT GGTGGGCCC ATCATAGGAG CTGTCTCGC TGGTGGCCTT 780
 TATGAGTATG TCTCTGTCC AGATGTTGAA TTCAAACGTC GTTTTAAAGA AGCCTTCAGC 840
 AAAGCTGCCC AGCAAAACAA AGGAAGCTAC ATGGAGGTGG AGGACACAGT GAGTCAGGTA 900
 GAGACGGATG ACCTGATTCT AAAACCTGGA GTGGTGCATG TGATTGACGT TGACCGGGGA 960
 GAGGAGAAGA AGGGGAAAGA CCAATCTGGA GAGGTATTGT CTTCAGTATG ACTAGAAGAT 1020
 CGCACTGAAA CGACACAAGA CTCCTTAGAA CTGCTCTCAG ATTTCCTTCC ACCCATTAAG 1080
 GAAACAGATT TGTATAAAT TAGAAATGTG CAGGTTGTGT GTTTCATGTC ATATTACTCA 1140
 GCTAAACAA TAAATATTTC ATAATTACA AAGGAGGAAC GGAAGAAACC TATTGTGAAT 1200
 TCCAAATCTA AAAAAGAAA TATTTTAAAG ATGTTCTTAA GCAATATAT ACCTATTTTA 1260
 TCTAGTTACC TTCTATTAC AACCAATTTT AACCGTGTGT CAAGATTGGG TTAAGTCTTG 1320
 CCGCAGAGAA CTCAAAGACA CGTCTATCAG CTTATTCTCT CTCTACTGGA ATATTGGTAT 1380
 AGTCAATTTT TATTTGAATA TTTATTCTAT TAACTGAGT TTAACAATGG C

Seq ID NO: 497 Protein sequence
 Protein Accession #: NP_001641.1

1 11 21 31 41 51
 MSDRPTARRW GKCGPLCTRE NIMVAFKQWV TQAFWKAVTA EFLAMLIFVL LSLGSTINWG 60
 GTEKPLPVDV VLISLCFGLS IATMVQCFFH ISGGHINPAV TVAMVCTRKI SIAKSVFYIA 120
 AQCLGAIIGA GILYLVTPPS VVGLGVTVH HGNLTAGHGL LVELIITFOL VFTIFASCDS 180
 KRTDVTGSIA LAIGFSVAIG HLFAINYTGA SMNPARSFGP AVIMGNWENH WIYVWGPPIG 240
 AVLAGGLYEV VFCPPVFEKR RFKEAFSKAA QQTGGSYMEV EDNRSQVETD DLILKPGVVH 300
 VIDVDRGEEK KGKDQSGEVL SSV

Seq ID NO: 498 DNA sequence
 Nucleic Acid Accession #: AB020684.1
 Coding sequence: 1..1744

1 11 21 31 41 51
 CCCCCTTGTC ATTAATACAT TAAAAAGATT CAATCTTTAC CCTGAGGTAA TTTTGGCCAG 60
 TTGGTACCGG ATTTATACCA AAATAATGGA CTTGATTGGT ATTCAAACCA AGATATGTTG 120
 GACGGTACCC AGAGGAGAAAG GACTCAGTCC TATTGAAAGC TGTGAAGGAT TGGGAGATCC 180
 TGCTTGCTTT TATGTTGCTG TAAATTTTAT TTTAAATGGA CTAATGATGG CATTATTCTT 240
 CATATATGGC ACATATTTAA GTACCCGTGT AATGTGACA CCACCTCTCC GTGAAAGCTT 300
 CTTTTCATAT CATGGAGAGT GTACCCGTGT AATGTGACA CCACCTCTCC GTGAAAGCTT 360
 CTCATATCCA TTCTTGTGTC TTCAGATGTT GCTAGTGACT CATATTCTCA GGGCTACAAA 420
 ACTTTATAGA GGAAGCTTGA TTGCACTCTG CATTTCCAAT GTATTTTCTA TGCTTCTCTG 480
 GCAGTTTGCT CAGTTTGTAC TTCTTACTCA GATTGCATCA TTATTGTCAG TATATGTTGT 540
 CGGGTACATT GATATATGTA AATTACGGA GATCATTTAT ATACATATGA TTTCTCTTGC 600
 ACTTTGTTTT GTTTGTATGT TTGGGAATC AATGTTATTA ACTTCTTAT ATGCTTCTTC 660
 TTTGTTAATT ATTTGGGGTA TTCTGGCAAT GAAACCATAT TTCTGAAAA TAAATGTATC 720

5
10
15
20
25
30
35
40
45
50
55
60
65

TGAACCTAGT TTATGGGTTA TTCAAGGATG TTTTGGGTTA TTTGGAACCTG TCATACCTTAA 780
ATACCTTGACA TCTAAAAATTT TTGGTATGTC AGATGAGCGCT CATATTGGCA ACTTACTAAC 840
ATCAAAATTC TTTAGTTTATA AGGATTTTGA TACTTTATTG TATACCTGTG CAGCGGAGTT 900
TGACTTTATG GAAAAAGAGA CTCCACTGAG ATACACAAAG ACATTATTGC TTCCAGTTGT 960
TCTTTAGTAG TTTGTGCTA TTGTTAGAAA GATTATTAGT GATATGTGGG GTTGCTTAGC 1020
TAAACACACAG ACACATGTAA GAAAAACCA GTTTGATCAT GGAGAGCTGG TTTACCATGC 1080
ATTGCAATTG TTAGCATATA CAGCCCTTGG TATTTTAATT ATGAGACTAA AACTCTTCTT 1140
GACACCACAC ATGTGTGTTA TGGCATCACT GATCTGCTCA AGACAGCTAT TTGGATGGCT 1200
CTTTTGCAAA GTACATCCTG GTGCTATTGT GTTTGCTATA TTAGCAGCAA TGTCAATACA 1260
AGGTTTCAGCA AATCTGCAAA CCCAGTGSAA TATTGTAGGG GAGTTTCAGCA ATTTGCCCCA 1320
AGAAGAACTT ATAGAATGGX TCAAAATATAG TACTAAACCA GATGCACTGT TTGCGGGTGC 1380
CATGCCACAG ATGGCAAGTG TTAAGCTCTC TGCACTTCGG CCCATTGTGA ATCATCCACA 1440
TTATGAAGAC GCAGGCTTGA GAGCCAGAAC AAAAATAGTA TACTCAATGT ATAGTCGGAA 1500
AGCAGCCGAA GAAGTGAAGC GAGAACTGAT AAAGTTAAAA GTGAACATAT ACATTCTAGA 1560
AGAGTCATGG TGTGTAAAGG GAAAAACTCC CTTATGTAAC CTCTTGGTGA AGGATTCCAA 1620
AGAAGATCCT GCCAATGCTG GGAAACTCC CTTATGTAAC CTCTTGGTGA AGGATTCCAA 1680
ACCTCACTTC ACCACTGTAT TCCAGAACAG TGTTTACAAA GTCCTAGAAG TTGTAAAGA 1740
ATGACTGCTA AGGACTGTGA TGGCATCACT TGCCTACGGA GAACTACATC TGTAACTGTT TTAATGTTTT 1800
GCTAAGTCAT GTGTGTTTCA TATCCCAAAA ACTTTTATAG GTAACCTTT TCAAAATAGAA 1860
AACGTTTTAT TTGTCGAATT TGAATGTCAT TCTAATTATA AAAATGACTT ACACCTTTAT 1920
CAATTGGTTA CTATTTCAAT GCACCCCTTA AAATTTGCTA TGCAATGAG TATATGCTTG 1980
TACTTGACTT TAATATTGT GCTAAAGTGA GCAAGCTAC CTGTATAAAG AAAACACAGT 2040
GGGTGTGAC AAGGATGACA TGAATAACA GGACAATTCT GACATGTAG GGGCTGATTT 2100
TATAGTGTAA GAATATTAA TGCCCTTGC TTTCTTTTC TGCCCTTGC TCTTGTCTTT 2160
TGGACATTC AGTGATTGA AGTTCTTCGG TCAATGTCAG CCCTGTCTAT AACTTGAGTT 2220
ACAGTAGATG GGGCAGACAT GGAGTGTGTT CTATATAAAA CTATCTGTTT GTTTTACTTC 2280
CTGTGCGCTT TTTGTTCTC TGTCTCTCTG TTAATGAAGC TTTTCTGCC CATTTATTAAT 2340
CCAACTCTCT GGACCTTGTG GTTAGGAAAT TCCCTTAACT TCCAGCCATA TGGCATTATC 2400
GTGTCTCTTT CTCTCTCTCT CTGTCTCTCT CTCTCTCTCT CTCTCCATA TTTTCTGTCA 2460
AATAAGTACT GTTTACTCAT TTAGTTGCTT ATCAAGTACT TATCTTGGT TTTAAAAAAA 2520
ATTAATGGTA ACTGTATTTT TCTCATTTT AGCATTTATC AAATGTTTAT ATTTTAAATAC 2580
CTTTAAACCA CTTTAAAGTT TTTTCATGTT TAATTTATAG TTTAAGAAAA ACTATTTTGA 2640
ACAAACCCAA ATATAGTGA TCTAGAAACT AATGTATATT TGATTAGACA TCATTTATAG 2700
TGGAACAGTA GACTGTAGTA CATGTAATT TTTCTTTTAC TATTAAGATA CAATAAAACA 2760
TGACTAATTT TGCTGTCAA AATGTAAAGA ATAATGATA ATGGAGTTT TTATATTTTA 2820
CTTTAAGAT TGCCGTCTCT TAATAAGACA AAGCCTTAAG CCTTATGTTA TAATTTTGGT 2880
TCTAAAAACC ATCATTTTCA TATAAGGAAT AAGTATATTT CGTCTCTCTC TTTAGTTTTT 2940
TTCTTCTAT TTATTTTAT TTGAAAAAT TTCTACACCT TCTTTGAATT CCTTGTATGA 3000
ATTTTGTGTT CTTAAGAGTT AATTGTGTG AAATGAGATT CTTCAAAACG ATGAAACCTC 3060
ATAGCTCTGA GAAAGGTTT TAGGGTTTTA AATCTAAGC AAAGCGTGAC TATGGCTGAC 3120
AGACTACACA TTTAATTATA CAGCTTCTCT TTTCTAACCA CAGGCAGATT AACCTCATTG 3180
TGGATTGTCC TTACAGCCTT AGTCTCTCAG CATGGTTTCT GGTGCCCATC CCGTGAAGCC 3240
GCTGTTCCTT TTTCTACCTC TTACCAAGAG CCAAGGGCAG CCCTGTGTCG GGGGAAGCAG 3300
CAGCTTGCTG ACATAAGTCA GCTGCAAGG CTGAGGAGTG TGCCCTCAGA GAAGCACCCG 3360
CCCCAGTCT TGTGCCAGC CTTAGAGCCG CAGCTCCAG GATGCTCCT TCCCTGGAGG 3420
CAGCCAGGA GAGGACTCT TCCAGAGTTG TGGCCACTGT TTCTCATTTG 3480
CTGGTTGACT GTTTTATTT CTTAGGCTTT TGCTAGTTTT AGAAAAAGG GAAGCAGCCC 3540
TTGATTGTG GATTAAAGC AACATTTGAG CGATGATGCA CAACAGTCCA GGAAAAATGG 3600
CGGTGGACAC TTGAGGCTGA GATGGGAGT TGACATGAGC AGGGAGAGGG AGGTGCGCGC 3660
TGCTTATCTG TGATTGTTGC TCACCTGAGT GTGGCTGATT GTGTACATCC AGCAGTTACA 3720
ATTTTAAAA ATTATACCTT TACATTTTAT TTATATTTTT CTCACCCCA GTAATTTCTT 3780
TCCAAAGAAG TTACATGTTA ATAAGTAGAA ATTCTGTATA GAAAAAAGC ATTTAAAAATA 3840
CTATATAAC TGCTTCATT GCTGGGAACC ATTAAGAGTA ATATAAATTA GCTTTTCCA 3900
GAAGGATCCT TTTGTAGCAG TGTATTATGA TGAACCCCA AGCAAAATAT GGCTATATAT 3960
TAGGGGAGCC AGTTTGGAGC AGAGGCCCTG AGSTCCCTGC TATGCAGCG TGGCCACAGC 4020
TGGCAGCCA AGCACTGTGG AGCATCCACA CCTTTGATGG CAATGCAGAT TGGTAGCAGG 4080
TTCCATAGGC GTACAAACA GTATTAAAGC TCAGTGTGTT GCATATTGTT AGCATTTACA 4140
AATATTGTTG CTTTAGTATG AGGAAAGTAA GSAATGGGCA AGAAGCGATC AAAATAGCTA 4200
TTGCTACAAC ATTTTCAAAA ACAAGTTGG GSGTGTATT CTTTAAAAAG ATAAGCCTCT 4260
AAAAATGCTT GGCAAAAAAA ATATAGTGTT AAAATAGGCC AGTGATATTA ATGAGAAAT 4320
GAAAGTATGT ATCAGGAATA AAGTGATATT GCATAGGAGT ATTGTATTTT TATGAATTTT 4380
ATGCCAGTTG TTTACATGTA CTATATATGT TAAATTAATA AAAATCATGA GAAATG

Seq ID NO: 499 Protein sequence
Protein Accession #: BAA74900.1

70
75
80
85

1 11 21 31 41 51
PLVINTLKRF NLYPEVILAS WYRIYTRIMD LIGIOTKICN TVTRGEGLSP IESCEGLGDP 60
ACFYVAVIFI ENGLMMALFF IYGYLSGSR LGGLVTVLCP FPNHGECTRV MWTPLRESF 120
SYFPLVLQML LVTHILRATK LYRGLIALC ISNVFFMLPW QFAQFVLLTQ IASLFAVYV 180
GYIDICKLRK IYIHMSLA LCFVLMFGNS MLLTSYASS LVIIWGLAM KPFLKINVS 240
ELSLWVIQGC FWLFGTVILK VLTSKIFGIA DDAHIGNLLT SKFFSVKDFD TLLVTCNAEF 300
DFMEKETPLR YTKTLLLPVV LVVFAIVRK IISDMWGLVA KQQTHVRKHQ FDHGLVYHA 360
LQLLAYTALG ILIMRLKPLF TPHMCVMASL ICSRQLFGWL FCKVHPGAIV FAILAAMSQ 420
GSANLQTQWN IVGEFNLQPO BELIEWIKYS TKPDVAFAGA MPTMASVKLS ALRPIVNHPH 480
YEDAGLRART KIVYSMSYRK AAEVVKRELI KLVNYIILE ESWCVRRSKP GCSMPETWDV 540
EDPANAGKTP LCNLLVDSK PHFTTVFQNS VYKLVLEVKE

Seq ID NO: 500 DNA sequence
Nucleic Acid Accession #: NM_001276.1
Coding sequence: 127..1278

85

1 11 21 31 41 51
AGTGGAGTGG GACAGGTATA TAAAGGAAGT ACAGGGCCTG GGGAAGAGGC CCTGTCTAGG 60
TAGCTGGCAC CAGGAGCCGT GGGCAAGGGA AGAGGCCACA CCCTGCCCTG CTCTGCTGCA 120

5
10
15
20
25
30

```

GCCAGAAATGG GTGTGAAGGC GTCTCAAACA GGCTTTGTGG TCCTGGTGCT GCTCCAGTGC 180
TGCTCTGCAT ACAAACTGGT CTGCTACTAC ACCAGCTGGT CCCAGTACCG GGAAGGCGAT 240
GGGAGCTGCT TCCGAGATGC CCTTGACCGC TTCTCTGTGA CCCACATCAT CTACAGCTTT 300
GCCAATATAA GCAACGATCA CATCGACACC TGGGAGTGGA ATGATGTGAC GCTCTACGGC 360
ATGTCTAACA CACTCAAGAA CAGGAACCCC AACCTGAAGA CTCTCTTGTC TGTCGGAGGA 420
TGGAACTTTG GGTCTCAAAG ATTTTCCAAG ATAGCCTCCA ACACCCAGAG TCGCGGACT 480
TTCATCAAGT CAGTACCGCC ATTCCTGCGC ACCCATGGCT TTGATGGGCT GGACCTTGCC 540
TGGCTCTACC CTGGACGGAG AGACAAACAG CATTTTACCA CCTAATCAA GGAAATGAAG 600
GCCGAATTTA TAAAGGAAGC CCAGCCAGGG AAAAAGCAGC TCCTGCTCAG CGCAGCACTG 660
TCTGCGGGGA AGGTCAACAT TGACAGCAGC TATGACATTG CCAAGATATC CCAACACCTG 720
GATTTTATTA GCATCATGAC CTACGATTTT CATGGAGCCT GGCCTGGGAC CACAGGCCAT 780
CACAGTCCCC TGTTCGAGG TCAGGAGGAT GCAAGTCTCT ACAGATTGAG CAACACTGAC 840
TATGCTGTGG GGTACATGTT GAGGCTGGGG GCTCCTGCCA GTAAGCTGGT GATGGGCATC 900
CCACCTTGG GGAGGAGCTT CACTCTGGCT TCTTCTGAGA CTGGTGTGG AGCCCCAATC 960
TCAGGACCGG GAATTCAGG CCGGTTCAAC AAGGAGGCGG GGACCTTGTC CTACTATGAG 1020
ATCTGTGACT TCTTCCGCGG AGCCACAGTC CATAGAACC TCGCCAGCA GGTCCCTAT 1080
GCCACCAAGG GCAACCAAGT GGTAGGATAC GACGACCAAG AAAGCGTCAA AAGCAAGGTG 1140
CAGTACTCTGA AGGATAGGCA GCTGGCAGGC GCCATGGTAT GGGCCCTGGA CTGGATGAC 1200
TTCAGGAGCT CCTTCTGCGG CCAGGATCTG CGCTTCCCTC TCACCAATGC CATCAAGGAT 1260
GCATCTGCTG CAACGTAGCC CTCTGTCTCT CACACAGCAC GGGGGCCAAAG GATGCCCGGT 1320
CCCCCTCTGC CTCTCAGCTG CCGGAGCCTT GATCACCTGC CCTGCTGAGT CCCAGGCTGA 1380
GCCTCAGTCT CCCCCTCTTG GGGCCTATGC AGAGGTCCAC AACACACAGA TTTGAGCTCA 1440
GCCCTGGTGG GCAGAGAGGT AGGGATGGGG CTGTGGGGAT AGTGAGGCAT CGCAATGTAA 1500
GACTCGGGAT TAGTACACAC TTGTTGATGA TTAATGGAAA TGTTCACAGA TCCCCAAGCC 1560
TGGCAAGGGA ATTTCTTCAA CTCCCTGCCC CCTAGCCCTC CTTATCAAAG GACACCATTT 1620
TGGCAAGCTG TATCAACCAAG GAGCCAAACA TCCTACAAGA CACAGTGACC ATACTAATTA 1680
TACCCCTGTC AAAGCCAGCT TGAACCTTC ACTTAGGAAC GTAATCGTGT CCCCTATCCT 1740
ACTTCCCTCT CTAATTTCCA CAGCTGCTCA ATAAAGTACA AGAGTTTAA AGTGTTGG 1800
CGCTTTGCTT TGGTCTATCT TTGAGCGCCC ACTAGACCCA CTGGACTCAC CTCGCCCATC 1860
TCTTCTGGGT TCCTTCCTCT GAGCCTTGGG ACCCTGAGC TTGCAGAGAT GAAGGCCGCG 1920
ATGTT

```

Seq ID NO: 501 Protein sequence
Protein Accession #: NP_001267.1

35
40
45

```

1 11 21 31 41 51
| | | | |
MGVKASQTGF VVLVLLQCCS AYKLVCYYS WSQYREGDGS CFPDALDRFL CTHIYSFAN 60
ISNDHIDTWE WNDVTLYGML NTLKNRNPNL KTLISVGGWN FGSQRFASKIA SNTQSRFFI 120
KSVPPFLRTH GFDGLDLAWL YPGRDRKHQF TLLIKEMKAE FIKEAQPGKK QLLLSAALSA 180
GKVTIDSSYD IAKISQHLDF ISIMTYDFHG AWRGTTGHHS PLFRGQEDAS PDRFSNTIDYA 240
VGYMLRLGAP ASKLVMGIPF FGSRFTLASS ETGVGAPISG PGIPGRFTKE AGTLAYYIEIC 300
DFLRGATVHR TLGQVQPYAT KGNQWVGYDD QESVKSQVQY LKDRQLAGAM VWALDLDDFO 360
GSFCQQLDRF PLTNAIKDAL AAT

```

Seq ID NO: 502 DNA sequence
Nucleic Acid Accession #: NM_006474.1
Coding sequence: 181..659

50
55
60
65

```

1 11 21 31 41 51
| | | | |
GCTGCTAGG GTCTGGAAG CTCGGGCACC CTCCTCTCC GGGGCTCCTG CTCCCACCCC 60
TCCGGCCCCC CCACGCTCGC GCTCCTCCAG GCTGGGCTG TGGCGCGGCT GCTTTAATT 120
TTCCCCCAGC TCAGAACTTT GCTGCTCGGC CCCCAGGAGA GCAACAACCT AACGGGAAG 180
ATGTGGAAGG TGTAGCTCT GCTCTCTGTT TTGGGAAGCG CGTCGCTCTG GGTCTGCGCA 240
GAAGGAGCCA GCACAGGCCA GCCAGAAGAT GACACTGAGA CTACAGGTTT GGAAGCGCGC 300
GTTGCCATGC CAGGTGCCGA AGATGATGTG GTGACTCCAG GAACCAAGCA AGACCGCTAT 360
AAGTCTGGCT TGACAACTCT GGTGGCAACA AGTGTCACCA GTGTACAGG CATTGCTATC 420
GAGGATCTGC CAACTTCAGA AAGCACAGTC CACGCGCAAG AACAAAGTCC AAGCGCCACA 480
GCTCAACAGC TGGCCACCAG TCACTCCACG GAGAAAGTGG ATGGAGACAC ACAGACAACA 540
GTTGAGAAAG ATGGTTTGT C AACAGTGACC CTGGTTGGAA TCATAGTTGG GGTCTTACTA 600
GCCATCGGTT TCATTGGTGG AATCATCGTT GTGGTTATGC GAAAATGTIC GGGGAAGGTAC 660
TCGCCCTAAA GAGCTGAAGG GTTACGCCCT GCTTGCCAAC GTGCTTTAAA AAAAGACCGT 720
TTCTGACTCT GTGGCCCTGT CCTGAGCTC GTGGGGAGAA GATGACCCTG GGAACATTG 780
CGGGCCCATC CAGATTCCAC GGTGACTTTC CGTTTGCCAA ATTAACCGAG GAAAGACCTT 840
TCACCAGATT TGGTTCTTAA ACTTT

```

Seq ID NO: 503 Protein sequence
Protein Accession #: NP_006465.1

70
75

```

1 11 21 31 41 51
| | | | |
MKVSAALLFV LGSASLWVLA EGASTGQPED DTETTGLEGG VAMPGAEDDV VTPGTSEDYR 60
KSLGLTLVAT SVNSVTGIRI EDLPTSESTV HAQEQSPTSAT ASNVATSHST EKVDGDTQTT 120
VERDGLSTVT LVGIIVGVLL AIGFIGGIIV VVMRKMSGRY SP

```

Seq ID NO: 504 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 62..895

80
85

```

1 11 21 31 41 51
| | | | |
CACTGCTCTG AGAATTTGTG AGCAGCCCTT AACAGGCTGT TACTTCACTA CAACTGACGA 60
TATGATCATC TTAATTTACT TATTTCTCTT GCTATGGGAA GACACTCAAG GATGGGGATT 120
CAAGGATGGA ATTTTTCATA ACTCCATATG GCTTGAACGA GCAGCCGGTG TGTACCACAG 180
AGAAGCACGG TCTGGCAAT ACAAGCTCAC CTACGCAGAA GCTAAGGCGG TGTGTGAATT 240
TGAAGGCGGC CATCTCGCAA CTTACAAGCA GCTAGAGGCA GCCAGAAAAA TTGATTTCAC 300

```


5
10
15
20

```
TGTCCTGTGCT GCTGGATGGA TGGCTAAGGG CAGAGTTGGA TACCCCATTT TGAAGCCAGG 360
GCCCAACTGTG GGATTTGGAA AAACCTGGCAT TATTGATTAT GGAATCCGTC TCAATAGGAG 420
TGAAAGATGG GATGCCATT GCTACAACCC ACAACGCAAG GAGTGTGGTG GCGCTCTTAC 480
AGATCCAAAG CAATTTTATA AATCTCCAGG CTTCCCAAT GAGTACGAAG ATAACCAAT 540
CTGCTACTGG CACATTAGAC TCAAGTATGG TCAGCGTATT CACCTGAGTT TTTTAGATTT 600
TGACCTTGAA GATGACCCAG GTTGCTTGGC TGATTATGTT GAAATATATG ACAGTTACGA 660
TGATGTCAT GCGCTTTGGG GAAGATACTG TGGAGATGAG CTTCCAGATG ACATCATCAG 720
TACAGGAAAT GTCATGACCT TGAAGTTTCT AAGTGTGCT TCACTGACAG CTGGAGGTTT 780
CCAAATCAAA TATGTTGCAA TGGATCCTGT ATCCAAATCC AGTCAAGGAA AAAATACAAG 840
TACTACTTCT ACTGGAATA AAAACTTTT AGCTGGAAGA TTTAGCCACT TATAAAAAAA 900
AAAAAAGGA TGATCAAAAC ACACAGTGT TATGTTGGAA TCTTTTGGAA CTCCTTTGAT 960
CTCACTGTTA TTATTAACAT TTATTTAATA TTTTCTAAA TGTGAAAGCA ATACATAATT 1020
TAGGGAATAT TGGAAATAT AGGAACTTT AAACGAGAAA ATGAAACCTC TCATAATCCC 1080
ACTGCATAGA AATAACAAGC GTTAACATTT TCATATTTT TTTCTTCAGT CATTTTTCTA 1140
TTTGTGTAT ATGTATATAT GTACCTATAT GTATTGTCAT TTGAAATTTT GGAATCCTGC 1200
TCTATGTACA GTTTGTATT ATACTTTTAA AATCTTGAAC TTTATAAACA TTTTCTGAAA 1260
TCATTGATTA TTTTACAAA ACATGATTT AAACAGCTGT AAAATATTCT ATGATATGAA 1320
TGTTTATGCT ATTATTTAAG CCTGTCTCTA TTGTTGGAAT TTCAGGTCAT TTTCATAAAT 1380
ATTGTTGCAA TAAATATCCT TGAACACACA AAAAAAAA AA
```

Seq ID NO: 505 Protein sequence
Protein Accession #: Eos sequence

25
30

```
1 11 21 31 41 51
| | | | |
MIILYLFLLL LWEDTQGWGF KDGIHNSIW LERAAGVYHR EARSQKYKLT YAEAKAVCEF 60
EGGHLATYKQ LEARKIGIFH VCAAGWMAKG RVGYPIVKPG PNCGFGKTGI IDYGIRLNRS 120
ERWDAYCYNP HAKCCGGVFT DPKQIFKSPG FPNYEDNQI CYWHIRLKYG QRIHLSFLDF 180
DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELFDIIIS TGNVMTLKFL SDASVTAGGF 240
QIKYVAMPDV SKSSQGNKNTS TTSTGNKNFL AGRFSLH
```

Seq ID NO: 506 DNA sequence
Nucleic Acid Accession #: NM_007115.1
Coding sequence: 69..902

35
40
45
50
55
60

```
1 11 21 31 41 51
| | | | |
GAATTCGCAC TGCTCTGAGA ATTTGTGAGC AGCCCTTAAC AGGCTGTAC TTTACTACAA 60
CTGACGATAT GATCATCTTA ATTTACTTAT TTCTCTTGCT ATGGGAAGAC ACTCAAGGAT 120
GGGGATTCAG GGATGGAAAT TTTCATAACT CCATATGGCT TGAACGAGCA GCCGGTGTGT 180
ACCAACAGAGA AGCAGCGTCT GGCAATACA AGCTCACCTA CGCAGAAGCT AAGGCGGTGT 240
GTGAATTTGA AGGCGGCCAT CTGCAACTT ACAAGCAGCT AGAGGCAGCC AGAAAAATTG 300
GATTTCATGT CTGTGCTGCT GGATGGATGG CTAAGGGCAG AGTTGGATAC CCCATTGTGA 360
AGCCAGGGCC CAATCATGTA TTTGGAAAAA CTGGCATTAT TGATTATGGA ATCCGCTCTCA 420
ATAGGAGTGA AAGATGGGAT GCCTATTGCT ACAACCCACA CGCAAGGAGG TGTGTTGGCG 480
TCTTTACAGA TCCAAGCGA ATTTTAAAT CTCCAGGCTT CCCAAATGAG TACGAAGATA 540
ACCAATCTG CTACTGGCAC ATTAGACTCA AGTATGGTCA GCGTATTAC CTGAGTTTCT 600
TAGATTTTGA CCTGAAGAT GACCCAGGTT GCTTGGCTGA TTATGTTGAA ATATATGACA 660
GTTACGATGA TGTCATGGC TTTGTGGGAA GATACCTGG AGATGAGCTT CCAGATGACA 720
TCATCAGTAC AGGAAATGTC ATGACCTTGA AGTTTCTAAG TGATGCTTCA GTGACAGCTG 780
GAGGTTTCCA AATCAATAT GTTGCAATGG ATCCTGTATC CAAATCCAGT CAAGGAAAAA 840
ATACAAGTAC TACTTCTACT GGAAATAAAA ACTTTTAGC TGGAAAGATT AGCCACTTAT 900
AAAAAATAAA AAGGATGATC AAAACACACA GTGTTTATGT TGGAACTCTT TGGAACTCCT 960
TTGATCTCAG TGTTATTATT AACATTATT TATTATTTT CTAAATGTGA AAGAAATACA 1020
TAATTTAGG AAGATGGGAA AATATAGGAA ACTTTAAAC AGAAAAATGA ACCTCTCATA 1080
ATCCCACTGC ATAGAAATAA CAAGCGTTAA CATTTCATA TTTTCTCTT TCAGTCAATT 1140
TTGATTTTGT GGTATAGTGA TATATGTACC TATATGTATT TGCAATTGAA ATTTTGAAT 1200
CCTGCTCTAT GTACAGTTTT GTATTATACT TTTTAAATCT TGAACCTTAT GAACATTTT 1260
TGAAATCATT GATTATTCTA CAAAAACATG ATTTTAAACA GCTGTAAAT ATTCTATGAT 1320
ATGAATGTTT TATGCAATT TTAAGCCTGT CTCTATTGTT GGAATTTTCA GTCAATTTCA 1380
TAAATATTGT TGCAATAAAT ATCCTTCGGA ATTC
```

Seq ID NO: 507 Protein sequence
Protein Accession #: NP_009046.1

65
70

```
1 11 21 31 41 51
| | | | |
MIILYLFLLL LWEDTQGWGF KDGIHNSIW LERAAGVYHR EARSQKYKLT YAEAKAVCEF 60
EGGHLATYKQ LEARKIGIFH VCAAGWMAKG RVGYPIVKPG PNXXFGKTGI IDYGIRLNRS 120
ERWDAYCYNP HAKCCGGVFT DPKRIFKSPG FPNYEDNQI CYWHIRLKYG QRIHLSFLDF 180
DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELFDIIIS TGNVMTLKFL SDASVTAGGF 240
QIKYVAMPDV SKSSQGNKNTS TTSTGNKNFL AGRFSLH
```

75
80

Seq ID NO: 508 DNA sequence
Nucleic Acid Accession #: NM_001044.1
Coding sequence: 129..1991

85

```
1 11 21 31 41 51
| | | | |
ACCGCTCCGG AGCGGGAGGG GAGGCTTCGC GGAACGCTCT CGGCGCCAGG ACTCGCGTGC 60
AAAGCCCAAG CCCGGGCGGC CAGACCAAGA GGAAGAAGC ACAGAATTC TCAACTCCCA 120
GTGTGCCCAT GAGTAAGAGC AAATGCTCCG TGGGACTCAT GTCTTCCGTG GTGGCCCGCG 180
CTAAGAGACC CAATGCCGTG GGGCCGAAGG AGGTGGAGCT CATCTTGTCT AAGGAGCAGA 240
ACGGAGTGCA GCTCACAGC TCCACCCTCA CCAACCGCGC GCAGAGCCCC GTGGAGGCC 300
AGGATCGGGA GACCTGGGGC AAGAAGATCG ACTTTCCTCT GTCGTCATT GGCTTTGCTG 360
TGGACCTGCG CAACGCTCGG CGGTTCCCTT ACCTGTGCTA CAAAAATGGT GCGGCTGCCT 420
```

5
10
15
20
25
30
35
40
45
50
55
60

TCCTGGTCCC CTACCTGCTC TTCATGGTCA TTGCTGGGAT GCCACTTTTC TACATGGAGC 480
TGGCCCTCGG CCAGTTCAAC AGGGAAGGGG CCGCTGGTGT CTGGAAGATC TGCCCCATAC 540
TGAAAGGTGT GGGCTTCAAG GTCATCTCTA TCTCACTGTA TGTGGGCTTC TTCTACAAAC 600
TCATCATCGC CTGGGCGCTG CACTATCTCT TCTCCTCCTT CACCAAGGAG CTCCCTCTGA 660
TCCACTGCAA CAACTCTCTG AACAGCCCCA ACTGCTCGGA TGCCATCTCT GGTGACTCCA 720
GTGGAGACAG CTCGGGCTCT AACAGACCTT TTGGGACCAC ACCTGCTGCC GAGTACTTTG 780
AACGTGGCGT GCTGCACCTC CACCAGAGCC ATGGCATCGA CGACCTGGGG CCTCCGCGST 840
GGCAGCTCAC AGCCTGCTCG GTGCTGGTCA TCGTGTCTCT CTACTTCAGC CTCTGGAAGG 900
GGGTGAAGAG CTCAGGGAAG GTGGTATGGA TCACAGCCAC CATGCCATAC GTGGTCTCTA 960
TGCCCTGCTC CCTGCGTGGG GTCACCTCTC CTGGAGCCAT AGACGGCATC AGAGCATACC 1020
TGAGCGTTGA CTCTACCGG CTCTGCGAGG CGTCTGTTTG GATTGACGCG GCCACCCAGG 1080
TGTGCTTCTC CTTGGGCTGT GGGTTGCGGG TGCTGATCGC CTCTCCAGC TACAACAAGT 1140
TCACCAACAA CTGCTACAGG GACGCGATTG TCACCACCTC CATCACTCC CTGACGAGCT 1200
TCTCCTCGG TCTGTCGCTC TTCTCCTTCC TGGGGTACAT GGCACAGAAG CACAGTGTGC 1260
CCATCGGGGA CGTGGCCAAG GACGGGCCAG GGCTGATCTT CATCATCTAC CCGGAAGCCA 1320
TGCCACGCTC CCTCTGTGCC TCAGCCTGGG CCGTGGTCTT CTTCATCATG CTGCTCACCC 1380
TGGGTATCGA CAGCGCCATG GGTGGTATGG AGTCAGTGAT CACCGGGCTC ATCGATGAGT 1440
TCCAGCTGCT GACCTACAGG CGTGAGCTCT TCACGCTCTT CATGCTCTCG GCGACCTTCC 1500
TCTGTCTCCT GTTCTGCGTC ACCAACGGTG GCATCTAAGT CTTCACGCTC CTGGACCAAT 1560
TTGCAGCCGG CACGTCCATC CTCTTTGGAG TGCTCATCGA AGCCATCGGA GTGGCTTGGT 1620
TCTATGTTGT TGGGCGTTGT AGGACGACA TCCAGCAGAT GACCGGGCAG CCGCCACGCC 1680
TGTACTGGCG GCTGTGCTGG AAGCTGGTCA GCCCTGCTTT TCTCTGTTT GTGGTCTGTTG 1740
TCAGCATGCT CAGCTTCAAG CCCCCCACT ACGGAGCCTA CATCTTCCC GACTGGGCCA 1800
ACGCGCTGGG CTGGGTCTAT GCCACATCTT CCATGGCCAT GGTGCCCATC TATGGGCGCT 1860
ACAAGTTCTG CAGCCTGCTT GGTCTCTTTC GAGAGAAACT GGCCTACGCC ATTGACCCG 1920
AGAAGGACCG TGAGCTGGTG GACAGAGGGG AGGTGCGCCA GTTCACGCTC CGCCACTGGC 1980
TCAAGGTGTA GAGGAGCAG AGACGAGAGC CCCAGGAAGT CATCTGCAA TGGGAGAGAC 2040
ACGAACAAAC CAGGAATAAT TAAGTTTCA GAGAAAGGAG GGCACTTCT ACTCTTCAAC 2100
CTCTACTGAA AACACAAACA ACAAGCAGA AGACTCCTCT CTCTGACTG TTTACACCTT 2160
TCCGTGCGGG GAGCGCACCT CGCGTGTCT TGTGTTGCTG TAATAACGAC GTAGATCTGT 2220
GCAGCGAGGT CCACCCCGTT GTTGTCTCTG CAGGGCAGAA AAACGTCTAA CTTCATGCTG 2280
TCTGTGTGAG GCTCCTCTCC TCCCTGCTCC CTGCTCCCG CTCTGAGGCT GCCCCAGGGG 2340
CACTGTGTTT TCAGGCGGGG ATCAGATCC TTGTAGAGCG ACCTGCTGAG AATCCCGGTG 2400
CTCACAGTAG CTCTCTAGAC CATTACTTT GCCCATATTA AAAAGCCAA GTCCTCTGTT 2460
GGTTTAGCTG TGCAGAAGT GAAATGGAGG AAACCACAA TTCATGCAA GTCTTTTCC 2520
GATGCGTGGC TCCAGCAGA GCGCGTAAAT TGAGGTTCA GTTGACAT TGCACACACA 2580
GTCTGTTGAG AGGCATTGGA GGATGGGGT CTTGTTATGT CTCACAGGA AATTCTGTTT 2640
ATGTTCTTGC AGCAGAGAGA AATAAACTC CTGAAACCA GCTCAGGCTA CTGCCACTCA 2700
GGCAGCTGTG GGTCTCTTGT GGTGTAGGGA ACGGCTGAG AGGAGCGTGT CCTATCCCC 2760
GACGATGCA GGGCCCCCAG AGGAGCGTGT CCTATCCCC GACGATGCA GGGCCCCCAG 2820
AGGAGCATGT CCTATCCTG GACGATGCA GGGCCCCCAG AGGAGCGTGT ACTACCCAG 2880
AACGATGCA GGGCCCCCAG AGGAGCGTGT ACTACCCAG GACGATGCA GGGCCCCCAG 2940
TGGAGCGTGT ACTACCCAG GACGATGCA GGGCCCCCAG AGGAGCGTGT CCTATCCCC 3000
GACCGAGCG ATGAGGGGCC CCCACAGGAG CGTGTACTAC CCCAGGAGCG ATGAGGGGCC 3060
CCCACAGGAG CGTGTACTAC CCCAGGATGC ATGAGGGGCC CCCACAGGAG CGTGTACTAC 3120
CCCAGGAGCG ATGAGGGGCC CCCATGCAAG CAGCCTGCG ACCAACACTC TGCCCTGGCT 3180
TGAGCCGTGA CCTCCAGGA GGGACCCAC TGGAAATTTA TTTCTCTCAG GTGCGTGCCA 3240
CATCAATAAC AACAGTTTT ATGTTTGGGA ATGGCTTTT AAAATCATAT TTACTGTGA 3300
ATCAAAACAA ATTCAAGAA GCAGTATCCG CGAGCCTGCT TGCTGATATT GCAGTTTTG 3360
TTTACAAGAA TAAITAGCAA TACTGAGTGA AGGATGTTG CCAAAGCTG CTTTCCATG 3420
CACACTGCCC TCTGCCACTG ACAGGAAAGT GGATGCCATA GTTTGAATTC ATGCTCTCA 3480
TCGGTGGGCC TGCTTACGTT CTGCCGAGG GCAGGGGGCG TGCAAGGCCA GTCTAGGCTG 3540
TCCCTGCAA GTGAGCTGG GCTCCAGGGA CTGAGTGTAT ATGCTCGGTG GGAGCCGTCA 3600
GCCTGTGAAC TGCCAGGCG CTGCAGTTAG CACAGAGGAT GGCTTCCCCA TTGCTTCTG 3660
GGGAGGGACA CAGAGGACG CTTCCCATC GCCTTCTGGC CGCTGCACTC AGCACAGAGA 3720
GCGGCTTCCC CATTGCCTTC TGGGAGGGA CACAGAGGAC AGTTTCCCCA TCGCTTCTG 3780
GTTGTGAAG ACAGACAGA GAGCGGCTTC CCCATCGCTC TCTGGGGAGG GGCTCCGTGT 3840
AGCAACCCAG GTGTTGTCG TGTCTGTTGA CCAATCTCTA TTCAGCATCG TGTGGGTCCC 3900
TAAGCACAAT AAAAGACATC CACAATGGAA AAAAAAAG GAATTC

Seq ID NO: 509 Protein sequence
Protein Accession #: NP_001035.1

65
70
75

1 11 21 31 41 51
MSKSKCSVGL MSSVAPAKE PNAVGPKEVE LILVKEQNGV QLTSSLTLP RQSPVEAQDR 60
ETWKKIDFL LSVIGFAVDL ANVWRFPYLC YKNGGGAFLV PYLLEFVIAG MPLPYMELAL 120
GQFNREGAAG WKICPILKG VGFTVILISL YVGFFYNVII AWALHYLFSS FTTELPIWIC 180
NNSWNSPNCs DAHPGDSGSG SSGLDNDFGT TPAEYFERG VLHLHQSHGI DDLGPWRQL 240
TACLVLVIVL LYFSLWKGVK TSGKVWITA TNPYVVLTA LRLGVTLPGA IDGIRAYLSV 300
DFYRLCEASV WIDAATQVCF SLGVGFGVLI APSSYNKFTN NCRYDAIVTT SINSLTSFSS 360
GFVVFSFLGY MAQKHSVPIG DVAKDGPGLI FIIYPEAIAT LPLSSAWAVV FFIIMLLTGI 420
DSAMGMEV ITGLIDEFQL LHRHRELFTL FIVLATFLLS LFCVTNGGIY VPTLLDHFPA 480
GTSILFGLVI EAIGVANFYG VGQFSDDIQ MTGQRPSLYW RLCWKLVSFC FLLPVVVVSI 540
VTFRPPHYGA YIFPDWALAN GWVIATSSMA MVPIYAAKFK CSLPGSPREK LAYAIAPKED 600
RELVDRCGEVR QFTLRHNLKV

Seq ID NO: 510 DNA sequence
Nucleic Acid Accession #: NM_001216.1
Coding sequence: 43..1422

80
85

1 11 21 31 41 51
GCCCCATAC ACCGTGTGCT GGGACACCCC ACAGTCAGCC GCATGGCTCC CCTGTGCCCC 60
AGCCCCCTGG TCCCTCTGTT GATCCCGGCC CCGTCTCCAG GCCTCACTGT GCAACTGCTG 120
CTGTCACTGC TGCTTCTGAT GCCTGTCCAT CCCCAGAGGT TGCCCCGGAT GCAGGAGGAT 180
TCCCCCTTGG GAGGAGGCTC TTCTGGGGAA GATGACCCAC TGGGCGAGGA GGATCTGCCC 240

5 AGTGAAGAGG ATTCACCCAG AGAGGAGGAT CCACCCCGAG AGGAGGATCT ACCTGGAGAG 300
 GAGGATCTAC CTGGAGAGGA GGATCTACCT GAAGTTAAGC CTAATCAGA AGAAGAGGGC 360
 TCCCTGAAGT TAGAGGATCT ACCTACTGTT GAGGCTCCTG GAGATCCTCA AGAACCCAG 420
 AATAATGCC ACAGGGACAA AGAAGGGGAT GACCAGAGTC ATTGGCCCTA TGGAGGCGAG 480
 CGCCCTTGCC CCGGGGTGTC CCCAGCCTGC GCGGGCCGCT TCCAGTCCCC GGTGGATATC 540
 CGCCCTCAGC TCGCCGCTCT CTGCCCGGCC CTGGCCCTCC TGGAACTCCT GGGCTTCCAG 600
 CTCCCGCCGC TCCAGAACT GCGCTGCGC AACAAATGCC ACAGTGTGCA ACTGACCCCTG 660
 CCTCTGGGC TAGAGATGGC TCTGGGTCCC GGGCGGAGT ACCTGCTCTG GCAGCTGCAT 720
 CTGCACTGGC GGGCTGCAGG TCGTCCGGGC TCGGAGCACA CTGTGAAGG CCACCGTTTC 780
 10 CCTGCCAGA TCCAGTGGT TCACCTCAGC ACCTGCTTGG CCAGAGTTGA CGAGGCCCTG 840
 GGGCGCCCGG GAGGCTGGC CGTGTGGCC GCCTTCTGG AGGAGGGCCC GGAAGAAAAC 900
 AGTGCCTATG AGCAGTGTCT GTCTCGCTTG GAAGAAATCG CTGAGGAAGG CTCAGAGACT 960
 CAGGTCCAG GACTGGACAT ATCTGCATC CTGCCCTCTG ACTTCAGCCG CTAATTCCAA 1020
 TATGAGGGGT CTCTGACTAC ACCGCCCTGT GCCCAGGGTG TCATCTGGAC TGTGTTTAA 1080
 15 CAGACAGTGA TGCTGAGTGC TAAGCAGCTC CACACCTCTT CTGACACCTT GTGGGGACCT 1140
 GGTGACTCTC GGCTACAGCT GAACTTCGGA GCGAGCGAGC CTTTGAATGG GCGAGTGATT 1200
 GAGGCCCTCT TCCCTGCTGG AGTGGACAGC AGTCTCGGG CTGCTGAGCC AGTCCAGCTG 1260
 AATTCCTGCC TGGCTGCTGG TGACATCTTA GCCCTGGTTT TTGGCTCCTT TTTTGTCTGC 1320
 20 ACCAGCGTGC CGTTCCTTGT GCAGATGAGA AGGCAGCACA GAAGGGGAAC CAAAGGGGGT 1380
 GTGAGTACC GCCCAGCAGA GGTAGCCGAG ACTGGAGCCT AGAGGCTGGA TCTTGGAGAA 1440
 TGTGAGAAGC CAGCCAGAGG CATCTGAGGG GGAGCCGCTA ACTGTCTGT CCGTCTCATT 1500
 ATGCCACTTC CTTTAACTG CCAAGAAATT TTTTAAATA AATATTTATA AT

Seq ID NO: 511 Protein sequence
 Protein Accession #: NP_001207.1

25 1 11 21 31 41 51
 MAPLCPSPWL PLLIPAPAPG LTVQLLSLL LMPVHPQRL PRMQEDSPLG GGSSGEDDPL 60
 30 GEEDLPSEED SPREDPPGE EDLPGEEDLP GEEDLPVKK KSEEGSLKL EDLPTVEAPG 120
 DPQSPQNNAH RKEGGDDQSH WRYGGDPPWP RVSPACAGRF QSPVDIRPQL AAFCPALRPL 180
 ELLGFQPPPL PELRLRNNGH SVQLTLPPGL EMALGPGRFY RALQLHLHWG AAGRPGSEHT 240
 VEGHRFPAEI KVVHLSTAFR RVDEALGRPG GLAVLAAFL EGPENSAYE QLSRLIEIA 300
 EEGSETQVPG LDISALLPSD FSRYPQYEGS LTPPCAGGV IWTVFQNTVM LSAKLHTLS 360
 35 DTLWGPDSR LQLNFRATQP LNRVIEASF PAGVSSPRA AEPVQLNSCL AAGDILALVF 420
 GLLFAVTSVA FLVQMRQRH RGTGGVSYR PAEVAETGA

Seq ID NO: 512 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..3978

40 1 11 21 31 41 51
 45 ATGGTGGGTG AAGGACCTTA CCTTATCTCA GATCTGGACC AGCGAGGCCG CGGAGATCC 60
 TTTCAGAAA GATATGACCC CAGCCTGAAG ACCATGATCC CAGTGCAGCC CTGTGCAAGG 120
 TTAGCACCCA ACCCGGTGGA TGATGCCGGG CTACTCTCCT TCGCCACATT TTCTGGCTC 180
 ACGCCGGTGA TGGTGAAGG CTACCGGCAA AGGCTGACCG TAGACACCTT GCCCCATTG 240
 TCGACATATG ACTCATCTGA CACCAATGCC AAAAGATTTC GAGTCTCTTG GGATGAAGAG 300
 50 GTAGCAAGGG TGGTCTCTGA GAAGGCTCT CTGAGCCACG TGGTGTGGA AITCCAGAGG 360
 ACACGCGTGT TGATGGACAT CGTGCCCAAC ATCTGTGCA TCATCATGGC AGCCATAGGG 420
 CCGACAGTTC TCATTACCCA AATCCTCCAG CAGACTGAGA GGACCTCTGG GAAAGTCTGG 480
 CTTGGCATTG GACTGTGCAT AGCCCTTTT GCCACGAGT TTACCAAGT CTTCTTTTGG 540
 GCCTTGGCTT GGGCCATCAA CTACCGCAGS GCCATCCGT TGAAGGTGGC GCTCTCCACC 600
 55 TTGGTTTTTG AAAACCTAGT GTCCCTCAAG ACATTGACCC ACATCTCTGT TGGCGAGGTG 660
 CTCAATATCA TGTCAGTGA TAGCTATTCT TTGTTGAAG CTGCCTTGT TTGCTCTTTG 720
 CCGAGCACA TCCGATCCT AATGGTCTT TGTCGGCGT ACCTCTTTT CATTCTGGGG 780
 CCCACAGCTC TCATCGGGAT ATCAGTGTAT GTCATATTCA TACCCGTCCA GATGTTTATG 840
 60 GCCAAGTCA ATTCAGCTTT CGAAGGTCA GCAATTTTGG TGACAGACAA GCGAGTTCAG 900
 ACAATGAATG AGTTCTGAC CTGCATCAGG CTGATCAAAA TGATATGCCG GGAGAAATCT 960
 TTTACCAACA CTATCCAAGA TATAAGAAGG AGGGAAGAA AATTACTGGA AAAAGCTGGA 1020
 TTTGTCCAAA GTGGAACTC TGCCCTGGCC CCCATCGTGT CCACCATAGC CATCGTGTG 1080
 ACATTATCCT GCCACATCCT CCTGAGACGC AAATCAGCG CACCCGTGGC ATTTAGTGTG 1140
 65 ATTGCCATGT TTAATGTAAT GAAGTTTTC ATTGCAATCT TGCCCTTCTC CATCAAGCA 1200
 ATGGCTGAAG CGAATGTCT TCTAAGGAGA ATGAAGAAA TTCTCATAGA TAAAGCCCC 1260
 CCATCTTACA TCACCCCAAC AGAAGACCCA GATACTGTCT TGCTTTTAGC AAATGCCACC 1320
 TTGACATGGG AGCATGAAG CAGCAGGAAA AGTACCCCAA AGAAATGCA GAACCAAGAA 1380
 AGGCATTAT GCAAGAAACA GAGGTGAGAG GCATACAGTG AGAGGAGTCC ACCAGCCAAG 1440
 GGAGCCACTG GCCCAGAGGA GCAAAGTGAC AGCCTCAAT CGGTCTGCA CAGCATAAGC 1500
 70 TTTGTGGTGA GAAAGTTATG TCGTTATCCC GAAGCCAGC TCCTGGCTTG GAGGTGGCCA 1560
 GCAGTGTGTT TTGGGAGAT CATCAGAGGA TACAGGCTC ATGGATTTTC TGCTAAAGAC 1620
 AAGGATGAAT CTAGAAGGCT TCTTACTTGG CCCCAAGAG TGGATAGGAC TCAAAGGGCA 1680
 GCCAATATAC TGGGGAAGAT CTTGGGAATA TGTGGGAATG TGGGAAGTGG AGTCAATGGA 1740
 75 ACTTTGGCCT ACGTTTACCA GCAGGCATGG ATCTTTTATG GAAATGTGAG AGAAACATA 1800
 CTCTTTGGAG AAAAGTATGA TCACCAAGG TATCAGACA CAGTCCGCTG CTGTGGCCTC 1860
 CAGAAGGACC TGACCAACCT CCCCATATGA GACCTGACTG AGATTGGGA GCGGGGCTC 1920
 AACCTCTCTG GGGGCGAGAG GCAGAGGATT AGCCTGGCCC GCGCTGTCTA CTCGACCGT 2040
 CAGCTCTACC TGCTGGACGA CCCCCTGTGC GCGTGGAGC CCCAGTGGG GAAGCACGTC 2100
 80 TTTGAGGAGT GCATTAAAGAA GACGCTCAGG GGAAGACAG TCGTCTGTGT GACCCACCAG 2160
 CTACAGTTCT TAGAGTCTTG TGATGAAGTT ATTTTATTAG AAGATGGAGA GATTGTGTA 2220
 AAGGGAACCC ACAAGAGGT AATGGAGGAG AGAGGGCGCT ATGCAAACT GATTCAACAC 2280
 CTGCGAGGAT TGCAGTTCAA GGATCCTGAA CACCTTACA ATGCAGCAAT GGTGGAAGCC 2340
 TTCAAGGAGA GCCCTGCTGA GAGAGAGGAA GATGCTGTA TAATCGGTA CTTCTTCT 2400
 85 CTTTCACTG TGTCTCTCTT CCTCTGATG ATTGGCAGCG CTGCTCTCAG CAACTGGTGG 2460
 CTGGGTCTCT GGTGGACAA GGGCTCAGG ATGACCTGTG GGGCCAGGG CAACAGGACC 2520
 ATGTGTGAGG TCGCGCGGT GCTGGCAGC ATCGGTGAGC ATGTGTACCA GTGGGTGTAC 2580
 ACTGCAAGCA TGGTGTTCAT GCTGCTGTTT GCGCTACCA AAGCTCTGT CTTACCAAG 2640

WO 02/086443

PCT/US02/12476

5
10
15
20
25

```

ACCACTGGA TGGCATCTC CTCTCTGCAT GACACGGTGT TTGATAAGAT CTTAAAGAGC 2700
CCAATGAGTT TCTTTGACAC GACTCCCACT GGCAGGCTAA TGAACCGTTT TTCCAAGGAT 2760
ATGGACGAGC TGGATGTGAG GCTGCCGTTT CACGCAGAGA ACTTTCTGCA GCAGTTTTTT 2820
ATGGTGGTGT TTAATCTCGT GATCTTGGCT GCTGTGTTTC CTGCTGTCTT TTTAGTCTGT 2880
GCCAGCCTTG CTGTAGGCTT CTTCATTCTG TTAGCGATTT TCCACAGAGG AGTCAGGAG 2940
CTCAGAAGG TGGAGAATGT CAGCGGTCA CCTGGTTCA CCCACATCAC CTCCTCCATG 3000
CAGGGCCTGG TGTCATTGTC ATACATCATC CAGCTGAGCG GACTGCTCCA AGTGTGTGTG 3060
TCCAAGGCC TGTCATTGTC ATACATCATC CAGCTGAGCG GACTGCTCCA AGTGTGTGTG 3120
CGAAGCGGAA CAGAGACGCA AGCCAAATTC ACCTCCGTGG AGTGTCTCAG GSAATACATT 3180
TGCACCTGTG TTCTGAATG CACTCATCCC CTCAAAGTGG GGACCTGTCC CAAGGACTGG 3240
CCCAGCTGTG GGGAGATCAC CTTGAGAGAC TATCAGATGA GATACAGAGA CAACACCCCC 3300
CTTGTCTCTG ACAGCCTGAA CTTGAACATA CAAAGTGGCG AGACAGTCGG GATTGTGGGA 3360
AGAACAGGTT CCGGAAAGTC ATCGTTAGGA ATGGCTTTGT TTCGTCTGGT GGAGCCAGCC 3420
AGTGGCACAA TCTTTATTGA TGAGGTGGAT ATCTGCATTC TCAGCTTGGG AGACCTCAGA 3480
ACCAAGCTGA CTGTAGTCCC ACAGGATCCT GTCCGTGTTG TAGGTACAGT AAGGTACAAC 3540
TTGGATCCCT TTSAGATGCA CACCGATGAG ATGCTCTGGC AGGTCTGAGA GAGAACATTC 3600
ATGAGAGACA CAATAATGAA ACTCCAGAA AAATACAGG CAGAAGTCAC AGAAAATGGA 3660
GAAAACCTCT CAGTAGGGGA ACGTCAGCTG CTTGTGTGGG CCCGAGCTCT TCTCGTAAT 3720
TCAAGATGCA TTCTCCTTGA TGAAGCCACC GCCTCTATGG ACTCCAAGAC TGACACCCCTG 3780
GTTGAGACA CCAACAAAGA TGCCCTCAAG GGCTGCACGT TGCTGACCAT CGCCCAACCGC 3840
CTCAACACAG TTCTCAACTC CGATCAGTCT CTGGTTATGG AAAATGGGAA GGTGATTGAG 3900
TTGACAAGC CTGAAGTCTT TGCAGAGAAG CCAGATTCTG CATTTCGAT GTTACTAGCA 3960
GCAGAAGTCA GATTGTAG

```

Seq ID NO: 513 Protein sequence
Protein Accession #: Eos sequence

30
35
40
45
50

```

1 11 21 31 41 51
MVGSGPYLIS DLDQRGRRRS FAERYDPSLK TMIFVRPCAR LAPNFVDDAG LLSFATFSWL 60
TPVMVKGYRQ RLTVDTLPLP STYDSSDTNA KRFRVLWDEE VARVGPEKAS LSHVVWKFQR 120
TRVLMDIVAN ILICIIMAAIG PTVLIHQILQ QTERTSQKVV VGIGLCIAIF ATEFTKVFVW 180
ALAWAINYRT AIRLKVALST LVFENLVSFK TLTHISVGEV LNILSSDSYS LFEAALFCPL 240
PATIPILMVF CAAYAFPILG PTALIGISVY VIPFVQMFPM AKLNSAFRRS AILVTDKRVQ 300
TMEFLTCIR LKMYANEKS FTNTIQDIRR RERKLEKAG FVQSGNSALA PIVSTIAIVL 360
TLSCHILLRR KLTAFAVAFSV IAMFNVMKFS IAILPFSIKA MAEANVSLRR MKKILIDKSP 420
PSYITQPEDP DTVLLANAT LTWEHEASRK STPKKLQNKQ RHLCKKQKSE AYSESPPAK 480
GATGFEEQSD SLKSVLHSIS FVVRKLCRYP EAQLLAWRWP AVFVGRIIRG YRPHGFSKAD 540
KDESRLTLTW PQEVDRTQRA AKYLGKILGI CGNVGSGRSS LLAALLGQMQ LQKGVVAVNG 600
TLAVVSQAW IFHGVNRENI LFGKEYDHQR YQHTVRVCGI QKDLNLPYG DLTEIGERGL 660
NLGGQQRRI SLARAVYSR QLYLLDDPLS AVDAHVGKIV FECEIKKTLR GKTIVLVTHQ 720
LQFLESCEDEV ILLEDGEICE KGTHELMEE RGRYAKLIHN LRGLQFKDPE HLYNAMVEA 780
FKEPFAEREE DAGIIGVLLS LFTVFLFLM IGSAAFSNW LGLWLDKGRS MTCGPQGNRT 840
MCEVGAVLAD IGQHYQWVY TASMVMLVF GVTGKGFVTK TTLMASSSLH DTVFDKILKS 900
PMSFDDTPT GRLMNRFSKD MDELDRVLPF HAENFLQOPF MVVFILVILA AVFPAVLLV 960
ASLAVGFIL LRIFHRGVQ LKKVENVSRS PWFTHITSSM QGLGIHAYG KKESCITYTS 1020
SKGLSLSYII QLSGLLQVCV RTGTETQAKF TSVELLREYI STCVPCTHP LKVGTCPKDW 1080
PSCGEITFRD YQMYRDNTP LVLDLNLNLT QSGQTVGIVG RTGSGKSLG MALFRLVEPA 1140
SGTIFIDEVD ICILSLEDLR TKLTVIPQDP VLFVGTVRYN LDPFESHTDE MLWQVLERTF 1200
MRDTIMKLEP KQAEVTENG ENFSVGERQL LCVARALLRN SKIILLDEAT ASDMSKTDTL 1260
VQNTIKDAFK GCTVLTIAHR LNTVLNCDHV LVMENGVIE FDKPEVLAEK PDSAFAMLLA 1320
AEVRL

```

Seq ID NO: 514 DNA sequence
Nucleic Acid Accession #: Z31560
Coding sequence: 1-966

60
65
70
75
80

```

1 11 21 31 41 51
CACAGCGCCC GCATGTACAA CATGATGGAG ACGGAGCTGA AGCCGCGGG CCCGCGACAA 60
ACTTCGGGGG CGCGCGGGCG CAACTCCACC CGCGGGCGG CCGGGCGGCA CCGAAAAAAC 120
AGCCCGGACC GGTCAAGCG GCCCATGAAT GCCTTCATGG TGTGGTCCC CGGGCAGCGG 180
CGCAAGATGG CCCAGGAGAA CCCCAAGATG CACAACCTCG AGATCAGCAA GCGCTTGGGC 240
GCGAGTGGAA AACTTTTGTC GGAGACGGAG AAGCGGCGGT TCATCGACGA GGCTAAGCGG 300
CTGCGAGCGC TGCATATGAA GGAGCACCGG GATTATAAAT ACCGCGCCCG CGCGAAAAAC 360
AAGACGCTCA TGAAGAAGGA TAAGTACACG CTGCGCGGCG GGCTGCTGGC CCGCGCGGCG 420
AATAGCATGG CGAGCGGGGT CGGGGTGGGC CGCGGCGTGG GCGCGGCGGT GAACCAAGCGC 480
ATGGACGATT ACGCGCATAT GAACGGCTGG AGCAACGGCA GCTACAGCAT GATGCAGGAC 540
CAGCTGGGCT ACCCGCAGCA CCGGGGCTTC AATGCGCAG CGCGAGCGCA GATGCAGCCC 600
ATGCACCGCT ACGAGTGAG CGCCCTGCAG TACAACCTCA TGACAGCTC GCAGACCTAC 660
ATGAACGGCT CGCCCACTTA CAGCATGTCC TACTGCGAGC AGGGCACCCC TGGCATGGCT 720
CTTGGCTCCA TGGGTTCTGT GGTCAAGTCC GAGGCGAGCT CCAGCCCCC TGTGGTTACC 780
TCTTCTCCCT ACTCCAGGGC GCCCTGCCAG CGCGGGGACC TCGGAGACAT GATCAGCATG 840
TATCTCCCGC GCGCGAGGT GCGGGAACCC GCGGCGCCCA GCAGACTTCA CATGTCCAG 900
CACTACCAGA GCGCGCGGT GCGCGGACCG GCCATTAAAG GCACACTGCC CCTCTCACAC 960
ATGTAGGGG CCGACAGCGA ACTGGAGGGG GGAGAAATTT TCAAAGAAAA ACGAGGAAAA 1020
TGGAGGGGT GCAAAAGAGG AGAGTAAGAA ACAGCATGGA GAAAACCGG TAGCTCAAA 1080
AAAA

```

Seq ID NO: 515 Protein sequence
Protein Accession #: CAA83435

85

```

1 11 21 31 41 51
HSARMYNNME TELKPPGPQO TSGGGGNGST AAAAGGNQKN SPDRVKRPMN AFMVWSRGQR 60
RRMAQENPM HNSEISKRLG AEWKLLSETE KRPFIDEAKR LRALHMKHP DYKYPRRRT 120
KTLMKDKYT LPGLLAPGG NSMASGVGVG AGLGAGVNR MDSYAHMNGW SNGSYSMMQD 180

```

QLGYPQHPGL NAHGAQMOP MERYDVSLQ YNSMTSSQTY MNGSPTYSMS YSQOQTPGMA 240
 LGSMSVVKKS EASSSPFVVT SSSHSRAPCO AGDLRDMISM YLPGAEPPEP AAPSRHMSQ 300
 HYQSGPVPPT AINGTLPLSH M

Seq ID NO: 516 DNA sequence
 Nucleic Acid Accession #: U91619
 Coding sequence: 29..541

1 11 21 31 41 51
 CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAAATCC AGCTTGTATG 60
 CATGCTACTC CTGGCTTTCA GCTCCTGGAG TCTGTGCTCA GATTGAGAAG AGGAAATGAA 120
 AGCATTAGAA GCAGATTCTT TGACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT 180
 TCCTCTTGG AAGATGACTC TGCTAAATGT TTGCAGTCTT GTAAATAATT TGAACAGCCC 240
 AGCTGAGGAA ACAGGAGAAG TTCATGAAGA GGAGCTTGTG GCAAGAAGGA AACTTCCTAC 300
 TGCTTAGAT GGCTTTAGCT TGGAAGCAAT GTTGACAATA TACCAGTCC ACAAAATCTG 360
 TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCTTG ATACTGGAAA 420
 TGACAAAAT GGAAAGGAAG AAGTCATAAA GAGAAAAATT CCTTATATTC TGAACCGGCA 480
 GCTGTATGAG AATAAACCCA GAAGACCCTA CATACTCAA AGAGATTCTT ACTATTACTG 540
 AGAGAATAAA TCATTTATTT ACATGTGATT GTGATTATC ATCCCTTAAT TAAATATCAA 600
 ATTATATTTG TGTGAAATG TGACAAACAC ACTTATCTGT CTCCTCTACA ATTGTGGTTT 660
 ATTGAATGTG TTTTTCGCA CTAATAGAAA TTAGACTAAG TGTTTTCAAA TAAATCTAAA 720
 TCTTCAAAA AAAAAAAAAA AAATGGGGCC GCAATT

Seq ID NO: 517 Protein sequence
 Protein Accession #: AAB50564

1 11 21 31 41 51
 MMAGMKIQLV CMLLLAFSSW SLCSDSSEEM KALEADFLTN MHTSKISKAH VPSWKMTLLN 60
 VCSLVNNLNS PAETETGEVHE EELVARRKLP TALDGFSLA MLTIYQLHKI CHSRAFOHWE 120
 LIQEDILDGT NDKNGKEEVI KRKIPYILKR QLYENKPRRP YILKRDSEYY

Seq ID NO: 518 DNA sequence
 Nucleic Acid Accession #: NM_006536.2
 Coding sequence: 109..2940

1 11 21 31 41 51
 ACCTAAAACC TTGCAAGTTC AGGAAGAACC CATCTGCATC CATAITGAAA ACCTGACACA 60
 ATGTATGCAG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACACAT GACCCAAAGG 120
 AGCATTGCAG GTCTATTGTC CAACCTGAAG TTTGTGACTC TCCTGGTTGC CTTAAGTTCA 180
 GAACCTCCAT TCCTGGGAGC TGGAGTACAG CTTCAGACA ATGGGTATAA TGGATTGCTC 240
 ATTGCAATTA ATCTCAGGT ACCTGAGAA CAGAACCTCA TCTCAACAT TAAGGAAATG 300
 ATAACGAAG CTTCATTTTA CCTATTTAAT GCTACCAAGA GAAGAGTATT TTTGAGAAAT 360
 ATAAAGATT TAATACCTGC CACATGGAAA GCTAATAATA ACAGCAAAAT AAAACAAGAA 420
 TCAATATGAA AGGCAAAATGT CATAGTGACT GACTGGTATG GGGCACATGG AGATGATCCA 480
 TACACCTCAC AATACAGAGG GTGTGGAAAA GAGGGAAAAA ACATTCAATT CACACCTAAT 540
 TTCCTACTGA ATGATAACTT AACAGCTGGC TACGGATCAC GAGGCCGAGT GTTTGTCCAT 600
 GAATGGGGCC ACCTCCGTTG GGGTGTGTTT GATGAGTATA ACAATGACA ACCTTCTCAC 660
 ATAAATGGGC AAAATCAAAAT TAAAGTGACA AGGTGTTTAT CTGACATCAC AGGCATTTTT 720
 GTGTGTGAAA AAGGTCTCTG CCCCCAAGAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA 780
 GGATGCACCT TTAATCAAAA TAGCACCCAA AATGCAACTG CATCAATAAT GTTCATGCAA 840
 AGTTTATCTT CTGTGGTTGA AITTTGTAAT GCAAGTACCC ACAACCAAGA AGCACCAAAAC 900
 CTACAGAAC AGATGTGCAG CCTCAGAAGT GCATGGGATG TAATCAGAGA CTCTGCTGAC 960
 TTTACACACA GCTTTCCCAT GAATGGGACT GAGCTTCCAC CTCCTCCAC ATTCCTGCTT 1020
 GTACAGGCTG GTGACAAAGT GGTCTGTTTA GTCTGGATG TGTCCAGCAA GATGGCAGAG 1080
 GCTGACAGAG TCGTCAACT ACAACAAGCC GCAGAAATTT ATTGATGCA GATTGTTGAA 1140
 ATTCATACCT TCGTGGCAT TGCCAGTTTC GACAGCAAG GAGAGATCAG AGCCAGCTA 1200
 CACCAAAATTA ACAGCAATGA TGATCGAAAG TTGCTGGTTT CATATCTGCC CACCACTGTA 1260
 TCAGCTAAAT CAGACATCAG CATTGTTTCA GGGCTTAAGA AAGGATTGTA GTGGTTGAA 1320
 AAACCTGAAT GAAAAGCTTA TGGCTCTGTG ATGATATTAG TGACCAGCGG AGATGATAAG 1380
 CTTCTGGCA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAATTCA CTCCATTGCC 1440
 CTGGGTTTAT CTGCGACCCC AAATCTGGAG GAATTATCAC GTCTTACAGG AGGTTTAAAG 1500
 TCTTTGTTTC CAGATATATC AAATCTCAAT AGCATGATTG ATGCTTTCAG TAGAATTTCC 1560
 TCTGGAAGTG GAGACATTTT CCAGCAACAT ATTCAGCTTG AAGTACAGG TGAAATGTC 1620
 AAAACCTCAC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT 1680
 ATGTTTCTAG TTAAGTGGA GGCAGTGGT CCTCTGAGA TTATATTATT TGATCTGAT 1740
 GGACGAAAT ACTACACAAA TAATTTATC ACCAATCTAA CTTTTCGGAC AGCTAGTCTT 1800
 TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCTGAAACAA TACCATCAT 1860
 TCTCTGCAAG CCTGAAAGT GACAGTGACC TCTGCGCCT CCACTCAGC TGTGCCCCCA 1920
 GCCACTGTGG AAGCCTTGTG GGAAAGAGAC AGCCTCCATT TTCCTCATCC TGTGATGATT 1980
 TATGCCAATG TGAACAGGG ATTTTATCCC ATTCTTAATG CCACTGTAC TGCCACAGTT 2040
 GAGCCAGAGA CTGAGATCC TGTTACGCTG AGACTCTTG ATGATGGAGC AGGTGCTGAT 2100
 GTTATAAAAA ATGATGGAAT TTAATCGAGG TATTTTCTCT CTTTGTCTGC AAATGGTAGA 2160
 TATAGCTTGA AAGTGATGT CAATCACTCT CCCAGCATAA GCACCCACAG CCACTCTATT 2220
 CCAGGGAGTC ATGCTATGTA TGTACCAGGT TACACAGCAA ACGGTAAAT TCAGATGAAT 2280
 GCTCCAAGGA AATCAGTAGG CAGAAATGAG GAGGAGCGAA AGTGGGGCTT TAGCCGAGTC 2340
 AGCTCAGGAG GCTCCTTTTC AGTCTGGGA GTTCCAGCTG GCCCCACCC TGATGTGTTT 2400
 CCACCATGCA AAATATTGA CCTGGAAGCT GTAAAAGTAG AAGAGGAATT GACCTATCT 2460
 TGGACAGCAC CTGGAGAAGA CTTTGTATCAG GGCAGGCTA CAAGCTATGA AATAAGAATG 2520
 AGTAAAGTC TACAGAAAT CCAAGATGAC TTTAACAATG CTATTTAGT AAATACATCA 2580
 AAGCGAAATC CTCAGCAAGC TGGCATCAGG GAGATATTTA OGTTCTCACC CCAGATTTC 2640
 ACGAATGGAC CTGAACATCA GCCAAATGGA GAAACACATG AAAGCCACAG AATTATGTT 2700
 GCAATAOCAG CAATGGATAG GAACTCCTTA CAGTCTGCTG TATCTAACAT TGCCAGGCG 2760
 CCTCTGTTTA TTCCCCCAA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAAA 2820
 GGAGTTTAA CAGCAATGGG TTTGATAGGA ATCAATTTGCC TTATTATAGT GTGACACAT 2880

5 CATACTTTAA GCAGGAAAAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATTATAA 2940
 ATAAATATCC AAAGTGTCTT CCTTCTTAGA TATAAGACCC ATGGCCTTCG ACTACAAAAA 3000
 CACTACTAAC AAGTCAAAAT AACATCAAAA CTGTATTAAA ATGCATTGAG TTTTGTGACA 3060
 ATACAGATAA GATTTTTTACA TGGTAGATCA ACAATTCTTT TGGGGGTAG ATTAGAAAAA 3120
 CCTTACACTT TGGCTATGAA CAAATAATAA AAATATTCTT TTAAGTAAT GTCTTTAAAG 3180
 GCAAGGGGAA GGCTAAAGTC GAACAGTGTG CAAGGAAAGT TTGTTTTATT GAGGTGGAAA 3240
 AATAGCCCCA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAACGTCTCG TGTGAAGCAA 3300
 TCATTAGTAT ACTTTGATTA ATTTTCTTT TCTCCTATC TGTGCAGTAC AGGTTGCTTG 3360
 10 TTTACATGAA GATCATGCTA TATTTTATAT ATGTAGCCCC TAATGCAAG CTCTTTACCT 3420
 CTGTCTATTT TGTATATAT ATTTTCTAGT ACATCTCCCT GCTAATGCTC AGAGATCTTT 3480
 TTTCACTGTA AGAGGTAACC TTTAACAATA TGGGTATTAC CTGTGTCTCT TCATACCGGT 3540
 TTTATGACAA AGGTCTATTG AATTATTATG TGTGTAAGTT TCTACTCCCA TCAAGCAGC 3600
 TTTCTAAGTT TATTGCCTTG GGTATTATG GAATGATAGT TATAGCCCCN TATAATGCCT 3660
 TACCTAGGAA A

Seq ID NO: 519 Protein sequence
 Protein Accession #: NP_006527.1

20 1 11 21 31 41 51
 MTQRSIAGPI CNLKFVTLV ALSSSELPFLG AGVQLQDNGY NGLLIAINPO VPENQNLISN 60
 IKEMITEASF YLFPNATKRRV PFRNIKILIP ATWKANNNSK IKQESYEKAN VIVTDWYGAAH 120
 GDDPYTLQYR CCGKEGKYIH FTFNELLNDN LTAGYGSRRGR VVHEWAHLR WGVDFEYNND 180
 25 KPFYINGQNO IKVTRCSSDI TGFVCEKGP CPQENCIISK LFKEGCTFIY NSTQNTASAI 240
 MFMQSLSSVV EFCNASTHNO EAPNLQNMOC SLRSADWVIT DSADFHHSFP MNGTELPPPP 300
 TFSLVQAGDK VVCLVLDVSS KMAEADRLLO LQQAEPFLM QIVRIHTFVG IASFDSKGET 360
 RAQLHQINSN DDKRLVSYL PTTVSAKTDI SICSLKKGF EVVEKLNGKA YGSVMILVTS 420
 GDDKLLGNCL PTVLSSGSTI HSIALGSSAA PNLEELSLRT GGLKFFVPDI SNSNSMIDAF 480
 30 SRISSGTGD I FQHQIQLEST GENVKPHQL KNTVTVDNTV QNDTMFLVTW QASGPPBIL 540
 FDPDGRKYT NNFITNLTFR TASLWIPGTA KPGHWYTYLN NTHSLQALK VVTVSRASNS 600
 AVPPATVEAF VVERSLHFFH PVMYIANVKQ GFYPILNATV TATVEPETGD PVTLRLLDDG 660
 AGADVINKDG IYSRYFFSFA ANGRYSKLVH VNHSPSISTP AHSIPGSHAM YVPGYTANGN 720
 IQMNAPRKSV GRNEEERKMG PSRVSSGGSP SVLGVPAGPH PDVFPCKII DLEAVKVEEE 780
 35 LTLNWTAPGE DFDGQATSY EIRMSKSLQN IQDDFNAIL VNTSKRNFPQ AGIREIPTFS 840
 PQISTNGPEH QFNGETHSH RIYVAIRAMD RNSLQSAVSN IAQAPLFIPP NSDPVPARDY 900
 LILKGVLTAM GLIGIICLII VVTHHTLSRK KRADKKENG T KLL

Seq ID NO: 520 DNA sequence
 Nucleic Acid Accession #: NM_000228.1
 Coding sequence: 82..3600

40 1 11 21 31 41 51
 45 GCTTTCAGGC GATCTGGAGA AAGAACGGCA GAACACACAG CAAGGAAGG TCCTTCTGCG 60
 GGATCACCCC ATTGGCTGAA GATGAGACCA TTCTTCTCTT TGTGTTTTCG CCTGCCTGCG 120
 CTCCTGCAATG CCCAACAGAG CTGCTCCCGT GGGGCTCGCT ATCCACCTGT TGGGACCTG 180
 CTTGTTGGGA GGACCCGGTT TCTCCGAGCT TCATCTACCT GTGGAAGTAC CAAGCCTGAG 240
 50 ACCTACTGCA CCCAGTATG CGAGTGGCAG ATGAAATGCT GCAAGTGTGA CTCACGAGCAG 300
 CCTCACAACT ACTACAGTCA CGAGTAGAG AATGTGGCTT CATCTCTCGG CCCCATGCGC 360
 TGGTGGCAGT CCCAGATGA TGTGAACCTT GTCTCTCTGC AGCTGGACCT GGACAGGAGA 420
 TTCCAGCTTC AAGAAGTCAT GATGGAGTTC CAGGGGCCCA TGCCCGCGGG CATGCTGATT 480
 GAGCGCTCCT CAGACTTCGG TAAGACCTGG CGAGTGTACC AGTACCTGGC TGCCGACTGC 540
 ACCTCCACCT TCCCTCGGGT CCGCCAGGGT CGGCCTCAGA GCTGGCAGGA TGTTCGGTGC 600
 55 CAGTCCCTCG CTCAGAGGCC TAATGCACGC CTAATGGGG GGAAGSTCCA ACTTAACTTT 660
 ATGGATTAG TGTCTGGGAT TCCAGCAACT CAAAGTCAAA AAATTCAGA GGTGGGGGAG 720
 ATCACAACTT TGTAGATCAA TTTACCCAGG CTGGCCCTCG TGCCCAAGG GGGCTACAC 780
 CCTCCAGCG CACTACTATG TGTGTCCAG CTCCGTCTGC AGGGGAGCTG CTTCTGTAC 840
 60 GGCCATGCTG ATCGCTCGGC ACCCAAGCCT GGGGCTCTG CAGGCCCTC CACCGCTGTG 900
 CAGGTCCAGG ATGTCTGTGT CTGCCAGCAC AACACTGCCG GCCCAAAATG TGAGCGCTGT 960
 GCACCCCTCT ACAACAACCG GCCCTGGAGA CCGGGGAGG GCCAGGAGC CCATGAATGC 1020
 CAAAGGTGCG ACTGCAATGG GCACTCAGAG ACATGTCACT TTGACCCCGG TGTGTTTGGC 1080
 GCCAGCCAGG GGGCATATGG AGGTGTGTGT GACAATTGCC GGGACCAAC CAAAGGCAAG 1140
 65 AACTGTGAGC GGTGTCACTT GCACTATTTC CGGAACCGGC GCCCGGAGC TTCCATTGAG 1200
 GAGACCTGCA TCTCTCGGGA GTGTGATCCG GATGGGGCAG TGCCAGGGGC TCCCTGTGAC 1260
 CCAGTGACCG GGCAGTGTGT GTGCAAGGAG CATGTGCAGG GAGAGCGCTG TGACCTATGC 1320
 AAGCCGGGCT TCACTGGACT CACCTACGCC AACCCGAGG GCTGCCACCG CTGTGACTGC 1380
 AACATCTGCG GGTCCCGGAG GACATGCCG GTGTACGAGG AGAGTGGCGG CTGCCTTTGT 1440
 70 CTGCCCAACG TGTGGGTCC CAAATGTGAC CAGTGTGCTC CTACCACTG GAAGCTGGCC 1500
 AGTGGCCAGG GCTGTGAACC GTGTGCTGCG GACCCGCACA ACTCCCTCA GCCACAGTG 1560
 CAACCACTTC ACAGGGCAGT GCCCTGTGCG GAAGGCTTTG GTGGCCTGAT GTGCAGCGCT 1620
 GCAGCCATCT GCCAGTGTCC AGACCGGACC TATGGAGACG TGGCCACAGG ATGCCGAGCC 1680
 TGTGACTGTG ATTTCCGGGG AACAGAGGGC CCGGCTGCG ACAAGGCATC AGGCGCTGC 1740
 75 CTCTGCGGCC CTGGCTTGAC CGGCOCGCG TGTGACCACT GCCAGCGAGG CTAAGTCAAT 1800
 CGCTACCCGG TGTGCGTGGC CTGCCACCCCT TGTCTCCAGA CCTATGATGC GGACCTCCGG 1860
 GAGCAGGCCC TGCGCTTTGG TAGACTCCCG AATGCCACCG CCAGCTGTG GTGAGGCGCT 1920
 GGGCTGGAGG ACCGTGGCCT GGCCTCCCGG ATCCTAGATG CAAAGAGTAA GATTGAGCAG 1980
 ATCCAGCAG TTTCTAGCAG CCCCAGATC ACAGAGCAGG AGGTGGCTCA GGTGGCCAGT 2040
 80 GCCATCTCTT CCTCAGGCG AACTCTCCAG GGCCTGCAGC TGGATCTGCC CCGTGGAGG 2100
 GAGACGTTGT CCTTCCGAG AGACCTGGAG AGTCTTGACA GAAGCTTCAA TGGTCTCTCT 2160
 ACTATGTATC AGAGGAAGAG GGAGCAGTTT GAAAAATAA GCAGTGCTGA TCCTTCAGGA 2220
 GCCTTCCGGA TGCTGAGCAC AGCCTACGAG CAGTCAGCCC AGGCTGCTCA GCAGGTCTCC 2280
 GACAGCTCAG GCCTTTTGA CAGCTCAGG GACAGCCGGA GAGAGGCAGA GAGGCTGGTG 2340
 CGGACGGCGG GAGGAGGAGG AGGCACCGGC AGCCCAAGC TGTGGCCCT GAGGCTGGAG 2400
 85 ATGCTTTCGT TGCTGACCT GACACCCACC TTCAACAAGC TGTGTGGCAA CTCACGGCAG 2460
 ATGGCTTGCA CCCCATAATC ATGCCCTGGT GAGCTATGTC CCAAGACAA TGGCAGCCG 2520
 TGTGCTCCC GGTGAGGGG TGTCTTCCC AGGGCCGGTG GGCCTTCTT GATGGCGGG 2580
 CAGGTGGCTG AGCAGCTGCG GGCCTTCAAT GCCAGCTCC AGCGACCAAG GCAGATGATT 2640

	AGGGCAGCCG	AGGAATCTGC	CTCACAGATT	CAATCCAGTG	CCCAGCGCTT	GGAGACCCAG	2700
	GTGAGCGCCA	GCCGCTCCCA	GATGGAGGAA	GATGTCAGAC	GCACACGGCT	CCTAATCCAG	2760
	CAGGTCCGGG	ACTTCTTAAC	AGACCCCGAC	ACTGATGCAG	CCACTATCCA	GGAGGTGAGC	2820
5	GAGGCGGTGC	TGGCCCTGTG	GCTGCCACCA	GACTCAGCTA	CTGTTCTGCA	GAAGATGAAT	2880
	GAGATCCAGG	CCATTGCAGC	CAGGCTCCCC	AACGTGGACT	TGGTGTGTCT	CCAGACCAAG	2940
	CAGGACATTG	CGCGTCCCCG	CCGGTTGCAG	GCTGAGGCTG	AGGAAGCCAG	GAGCCGAGCC	3000
	CATGCAGTGG	AGGGCCAGGT	GGAAGATGTG	GTGGAAGAAC	TGCGGCAGGG	GACAGTGGCA	3060
	CTGCAGGAAG	CTCAGGACAC	CATGCAAGGC	ACCAGCCGCT	CCCTTCGGCT	TATCCAGGAC	3120
	AGGGTTGCTG	AGGTTTCAGCA	GGTACTGCGG	CCAGCAGAAA	AGCTGGTGAC	AAGCATGACC	3180
10	AAGCAGCTGG	GTGACTTCTG	GACACGGATG	GAGGAGCTCC	GCCACCAAGC	CCGGCAGCAG	3240
	GGGGCAGAGG	CAGTCCAGGC	CCAGCAGCTT	GCGGAAGGTG	CCAGCGAGCA	GGCATTGAGT	3300
	GCCCAAGAGG	GATTITGAGAG	AATAAAACAA	AAGTATGCTG	AGTTGAAGGA	CCGTTTGGGT	3360
	CAGAGTTCCA	TGCTGGGTGA	GCAGGGTGCC	CGGATCCAGA	GTGTGAAGAC	AGAGGCAGAG	3420
	GAGCTGTTTG	GGGAGACCAT	GGAGATGATG	GACAGGATGA	AAGACATGGA	GTGAGAGCTG	3480
15	CTGCGGGGCA	CCAGGGCCAT	CATGCTGCGC	TGCGGGGACC	TGACAGGACT	GGAGAAGCGT	3540
	GTGGAGCAGA	TCCGTGACCA	CATCAATGGG	CGCGTGCTCT	ACTATGCCAC	CTGCAAGTGA	3600
	TGCTACAGCT	TCCAGCCCGT	TGCCCACTC	ATCTGCCGCC	TTTGCTTTTG	GTGSGGGGCA	3660
	GATTGGGTGG	GAATGCTTTC	CATCTCCAGG	AGACTTTCAT	GCAGCCTAAA	GTACAGCCTG	3720
	GACCAACCCCT	GGTGTGTAGC	TAGTAAGATT	ACCCTGAGCT	GCAGCTGAGC	CTGAGCCAAAT	3780
20	GGGACAGTTA	CACITGACAG	ACAAAGATGG	TGGAGATTGG	CATGCCATTG	AAACTAAGAG	3840
	CTCTCAAGTC	AAGGAAGCTG	GGCTGGGCAG	TATCCCCCGC	CTTTAGTTCT	CCACTGGGGA	3900
	GGAATCTCTG	ACCAAGCACA	AAAACCTAAC	AAAAGTGATG	TAAAAATGAA	AAGCCAAATA	3960
	AAAATCTTTG	G					

25 Seq ID NO: 521 Protein sequence
Protein Accession #: NP_000219.1

	1	11	21	31	41	51	
30	MRPFLLCFA	LEPLLHAQQA	CSRGACYPPV	GDLLVGRTRF	LRASSTCGLT	KPETYCTQYG	60
	EWQMKCKCD	SRQPHNYSH	RVENVASSSG	PMRWQSQND	VNPFVSLQLD	DRRFQLQEVN	120
	MEFQGPMPAG	MLIERSSDFG	KTWVRYQYLA	ADCTSTFPRV	RQGRPQSWQD	VRQSLPQRP	180
	NARLNGGKVQ	LNLMDLVSGT	PATQSQKIQE	VGEITNLRVN	FTRLAPVQPR	GYPHPSAYYA	240
	VSQRLQGGSC	FCHGHADRC	PKPGASAGPS	TAVQVHDVCV	CQHNTAGPNC	ERCAFYNNR	300
35	PWRPAEQDA	HECQRDCDNG	HSETCHFDPA	VFAASQGAYG	GVCNCRDHT	EGKNCERCQL	360
	HYFRNRPPGA	SIQETCISCE	CDPDGAVPGA	PCDPVTGQCV	CKEHVQGERC	DLCKPFTGL	420
	TYANPQCHHR	CDCNILGSR	DMPDDESGR	CLCLPNVVG	KCDQCAPYHW	KLASSQGCPE	480
	CACDPHNSPQ	PTVQPVHRAV	PCREGFGLM	CSAAAIRQCP	DRTYGDVATG	CRACDCDFRG	540
	TEGPGCDKAS	GRCLCRPGLT	GPRCDQCORG	YCNRYPCVVA	CHPCFQTYDA	DLREQALRFG	600
40	TLQGLQLDLP	LEEEETLSLR	DLESIDRSFN	GLLTMVQRKR	EQFEKISSAD	PSGAFRMLST	660
	AYEQSAQAQ	QVSDSSRLLD	QLRDSRREAE	RLVRQAGGGG	GTGSPKLVAL	RLEMSSLPDL	720
	TPTFNKLCGN	SRQMACTPIS	CPGELCPQDN	GTACGSRCRG	VLPRAAGAFI	MAGQVABQLR	780
	GFNAQLQRT	QMIRAAEESA	SQIQSSAQLR	ETQVSASRSQ	MEEDVRRTRL	LIQQVRDFLT	840
45	DPDIDAATIQ	HEVSEAVLAW	LPTDSATVLQ	KMEIQALAA	RLENVDLVLS	QTKQDIARAR	900
	RLQAEAEER	SRHAHEGQV	EDVVGNLRQG	TVALQEAQDT	MQGTSRSLRL	IQDRVAEVOQ	960
	VLRPAEKLV	SMTKQLGDFW	TRMEELRHQA	RQQGAEEVQA	QQLAEGASEQ	ALSAQEGFER	1020
	IKQKYAELKD	RLGQSSMLGE	QGARIQSVKT	EAEELFGETM	EMMDRMKDME	LELLRGSQAI	1080
	MLRSADLTGL	EKRVEQIRDH	INGRVLYYAT	CK			1140

50 Seq ID NO: 522 DNA sequence
Nucleic Acid Accession #: NM_001944.1
Coding sequence: 84..3083

	1	11	21	31	41	51	
55	TTTCTTAGA	CATTAACATGC	AGACGGCTGG	CAGGATAGAA	GCAGCGGCTC	ACTTGGACTT	60
	TTTCAACAGG	GAATTCAGAG	ACAATGATGG	GGCTCTTCCC	CAGAACTACA	GGGGCTCTGG	120
	CCATCTTCGT	GGTGGTCATA	TTGGTTTCATG	GAGAATTGCG	AATAGAGACT	AAAGGTCAAT	180
60	ATGATGAAGA	GAGATGACT	ATGCAACAAG	CTAAAGAAG	GCAAAAACGT	GAATGGGTGA	240
	AATTTGCCAA	ACCTGTCAGA	GAAGGAGAAG	ATAACTCAAA	AAGAAACCCA	ATTGCCAAGA	300
	TTACTTCAGA	TTACCAAGCA	ACCCAGAAAA	TCACCTACCG	AATCTCTGGA	GTGGGAATCG	360
	ATCAGCCGCC	TTTGGAAATC	TTTGTGTGTG	ACAAAAACAC	TGGAGATATT	AACATAACAG	420
	CTATAGTCGA	CCGGAGGAAA	ACTCCAAGCT	TCCTGATCAC	ATGTCGGGCT	CTAAATGCCC	480
65	AAGGACTAGA	TGTAGAGAAA	CCACTTATAC	TAACGGTTAA	AATTTTGGAT	ATTAATGATA	540
	ATCCTCCAGT	ATTTTCAACA	CAAAATTTCA	TGGGTGAAAT	TGAAGAAAAT	AGTGCCCTCA	600
	ACTCACTGGT	GATGATACTA	AATGCCACAG	ATGCAGATGA	ACCAAAACCA	TTGAATTCTA	660
	AAATTGCTCT	CAAAATTGTC	TCTCAGGAAC	CAGCAGGCAC	ACCCATGTTT	CTCCTAAGCA	720
	GAAACACTGG	GGAAAGTCCG	ACTTTGACCA	ATTCTCTTGA	CCGAGAGCAA	GCTAGCAGCT	780
70	ATCGTCTGGT	TGTGAGTGGT	GCAGACAAAG	ATGGAGAAGG	ACTATCAACT	CAATGTGAAT	840
	GTAATATTAA	AGTGAAAGAT	GTCAACGATA	ACTTCCCAAT	GTTTAGAGAC	TCTCAGTATT	900
	CAGCAOGTAT	TGAAGAAAAA	ATTTTAAGTT	CTGAATTACT	TCGATTTCAA	GTAACAGATT	960
	TGGATGAAGA	GTACACAGAT	AATTGGCTTG	CAGTATATTT	CTTTACCTCT	GGGAATGAAG	1020
	GAAATTTGGT	TGAATACAAA	ACTGATCCTA	GAACATAATG	AGGCATCCTG	AAAGTGGTGA	1080
75	AGGCTCTAGA	TTATGAACAA	CTACAAAGCG	TGAACCTTAG	TATTGCTGTC	AAAAACAAAG	1140
	CTGAATTTCA	CCAATCAGTT	ATCTCTCGAT	ACCGAGTTCA	GTCAACCCCA	GTCACAATTC	1200
	AGGTAATAAA	TGTAAGAGAA	GGAATTGCAT	TCCGTCCTGC	TTCCAAGACA	TTTACTGTGC	1260
	AAAAAGGCA	AAGTAGCAAA	AAATTGGTGG	ATTATATCCT	GGGAACATAT	CAAGCCATCG	1320
	ATGAGGACAC	TAAACAAGCT	CGCTCAAATG	TCAAATATGT	CATGGGACGT	AACGATGGTG	1380
80	GATACCTAAT	TGATGATTCA	AAAACGCTG	AAATCAAAAT	TGTCAAAAT	ATGAACCGAG	1440
	ATTCTACTTT	CATAGTTAAT	AAAACAATCA	CAGCTGAGGT	TCTGGCCATA	GATGAATACA	1500
	CGGGTAAATC	TTCTACAGGC	ACGGTATATG	TGAGAGTACC	CGAATTCAT	GACAATTTCT	1560
	CAACAGCTGT	CCTCGAAAAA	GATGCAGTTT	GCAGTTCTTC	ACCTTCGGTG	GTGTTCTCCG	1620
	CTAGAACACT	GAATAATAGA	TACACTGGCC	CCTATACATT	TGCACTGGAA	GATCAACCTG	1680
85	TAAAGTTGCC	TGCCGTATGG	AGTATCAAA	CCCTCAATGC	TACCTCGGCC	CTCCTCAGAG	1740
	CCAGGAACA	GATACCTCCT	GGAGTATACC	ACATCTCCCT	GGTACTTACA	GACAGTCAGA	1800
	ACAATCGGTG	TGAGATGCCA	CGCAGCTTGA	CAGTGGAGGT	CTGTCAGTGT	GACAACAGGG	1860

5
10
15
20
25

GCATCTGTGG AACTTCTTAC CCAACCACAA GCCCTGGGAC CAGGTATGGC AGGCGCACT 1920
CAGGGAGGCT GGGGCGTGGC GCCATGGGCC TGCTGCTCTT TGGTCTCTGT CTGCTGCTGT 1980
TGGCCCCCTT TCTGCTGTGT AACTGTGACT GTGGGGCAGG TTCTACTGGG GGAGTGACAG 2040
GTGGTTTAT CCCAGTTCTT GATGGCTCAG AAGGAACAA TCATCAGTGG GGAATTGAAG 2100
GAGCCCATCC TGAAGACAAG GAAATCACAA ATATTGTGT GCCTCTGTGA ACAGCCAATG 2160
GAGCGATTT CATGAAAAGT TCTGAAGTTT GTACAAATAC GTATGCCAGA GGCACAGCGG 2220
TGGAAAGCAC TTCAGGAATG GAAATGACCA CTAAGCTTGG AGCAGCCACT GAATCTGGAG 2280
GTGCTGCAGG CTTTGCACAA GGGACAGTGT CAGGAGCTGC TTCAGGATTC GGAGCAGCCA 2340
CTGGAGTTGG CATCTGTCTC TCAGGGCAGT CTGGAACCAT GAGAACAAAG CATTCCACTG 2400
GAGGAACCAA TAAGACTTAC GCTGATGGGG CGATAAGCAT GAATTTCTGT GACTCCTACT 2460
TTTCTCAGAA AGCATTTGCC TGTGCGGAGG AAGACGATGG CCAGGAAGCA AATGACTGCT 2520
TGTGATCTA TGATAATGAA GCGCGAGATG CCACTGGTTC TCTGTGGGCG TCCGTGGGTT 2580
GTTGCAGTTT TATTGCTGAT GACCTGGATG ACAGCTTCTT GGACTCAGTT GGACCCAAAT 2640
TTAAAAAAC TGCAGAGATA AGCCTTGGTG TTGATGGTGA AGGCAAGAA GTTCAGCCAC 2700
CCTCTAAGA CAGCGGTAT GGGATTGAAT CCTGTGGCCA TCCATAGAA GTCCAGCAGA 2760
CAGGATTGT TAAGTGCCAG ACTTTGTGAG GAAGTCAAG AGCTTCTGT TTGTCCGCTT 2820
CTGGGTCTGT CCAGCCAGCT GTTCCATCC CTGACCTCT GCAGCATGGT AACTATTAG 2880
TAACGGAGG TTACTCGCT TCTGTTTCCC TCTGCAACC TTCCACTGCA GGCTTTGATC 2940
CACTTCTCAC ACAAAATGT ATAGTGACAG AAAGGATGAT CTGTCCCAT TCCAGTGTTC 3000
CTGGCAACCT AGCTGGCCCA ACGCAGCTAC GAGGGTCACA TACTATGCTC TGACAGAGG 3060
ATCCTTGCTC CGTCTAATA TGACCAGAA GAGCTGGAAT ACCACATGA CCAATCTGG 3120
ATCTTTGGAC TAAAGTATTC AAAATAGCAT AGCAAGCTC ACTGTATTGG GCTAATAATT 3180
TGGCACTTAT TAGCTTCTCT CATAAATGA TCACGATTAT AAATTAATG TTTGGGTCA 3240
TACCCCAAAA GCAATATGTT GTCACTCTTA ATTCTCAAGT ACTATTCAA TTGTAGTAAA 3300
TCTTAAAGT TTTCAAACCT CTAATCAT ATTCG

Seq ID NO: 523 Protein sequence
Protein Accession #: NP_001935.1

30
35
40
45
50

1 11 21 31 41 51
MMGLFPRTTG ALAIFVVVIL VHGLRIETK QYDEEEMTM QAKRRQKRE WVKFAKPCRE 60
GEDNSKRNP: AKITSDYQAT QKITYRISGV GIDQPPFGIF VVDKNTGDIN ITAIVDREET 120
PSFLITCRAL NAQGLDVEKP LILTVKILD: NDNPVVFSSQ IFMGEIENS ASNSLVMLN: 180
ATDADEPNHL NSKIAFKIVS QEPAGTPEML LSRNTGEVET LTNLDREQA SSYRLVVSQA 240
DKDGEGLSTQ CECNKKVQDV NDNPMPRDS QYSARIEENI LSSELLRFQV TDLDEEYTDN 300
WLAVYFFTSN NEGNWFETQ: DPTNEGILK VVKALDYBQL QSVKLSIAVK NKAEFHQSVI 360
SRYRVQSTPV TIQVINREG IAFRPAKTF: TVQKGISSKK LVDYILGTQY AIDEDTNKAA 420
SNVKVVMGRN DGGYLMIDSK TAEIKFVKNM NRDSTFIVNK TITAEVLAI: EYTGKTSTGT 480
VYVRVPDFND NCPTAVLEKD AVCS SSPSV: VSARTLNNRY TPGYTPALED QPVKLPVWS 540
ITTLNATSAL LRAQEQIPFG VYHISLVLD: SQNNRCEMPR SLTLEVCQCD NRGICGTSYP 600
TTSPGTRYGR PHSGRLGPAA: IGLLLGLLL LLLAPLLLT CDGAGSTGG VTGGFIPVD 660
GSEGTIHQWG IEGAHPEDEK: ITNICVPVPT ANGADFMESS EVCTNTYARG TAVEGTSME 720
MTTKLGAATE SGGAGFATG TVSGAASGFG AATGVGICSS QSGTMRTRH STGGTNKDYA 780
DGAISMNPLD SYFSQKAFAC AEEDDQGEAN DCLLIYDNEG ADATGSPVGS VGCCSFIADD 840
LDDSPDLSL PKFKKLAIS LGVDGEGKEV QPPSKDSYGV IESCGHP:IEV QQTGFVKCQT 900
LSGSQGSAL SASGSVQPAV SIPDPLQGN LVTETYSAS GSLVQPSAG FDPILLQNV 960
VTERVICPIS SVFENLAGPT QLGRSHTMLC TEDPCSRLL

Seq ID NO: 524 DNA sequence
Nucleic Acid Accession #: XM_058069.2
Coding sequence: 1..1413

55
60
65
70
75
80

1 11 21 31 41 51
ATGAAGTTTC TTCTAATACT GCTCCTGCAG GCCACTGCTT CTGGAGCTCT TCCCTGAAC 60
AGCTCTACAA GCCTGGAAAA AATAATGTG CTATTTGGTG AAAGATACTT AGAAAAATT 120
TATGGCCTTG AGATAAACAA ACTTCCAGTG AAAAAATGA AATATAGTGG AAACCTAATG 180
AAGGAAAAAA TCAAGAAAA GCAGCACTTC TTGGGTCTGA AAGTGACCGG GCAACTGGAC 240
ACATCTACCC TGGAGATGAT GCAGCACTTC CGATGTGGAG TCCCGATGT CCATCATTT 300
AGGGAATGTC CAGGGGGGCC CGTATGGAGG AAACATTATA TCACCTACAG AATCAATAAT 360
TACACACCTG ACATGAACCG TGAGGATGTT GACTACGCAA TCCGAAAGC TTTCCAAGTA 420
TGGAGTAATG TTACCCCTCT GAAATTCAGC AAGATTACA CAGGCATGSC TGACATTTTG 480
GTGGTTTTTG CCGGTGGAGC TCATGGAGAC TTCCATGCTT TTGATGGCAA AGGTGGAATC 540
CTAGCCCATG CTTTTGAGC TGGATCTGCG ATTGGAGGGG ATGCACATT CGATGAGGAC 600
GAATTCCTGA CTACACATTC AGGAGGCACA AACTTGTTC TCACCTGCTG TCACGAGATT 660
GGCCATTCTT TAGGTCTTGG CCATTCTAGT GATCCAAAG CCGTAATGTT CCCCACCTAC 720
AAATATGTTG ACATCAACAC ATTTGCCTCT TCTGCTGAT ACATACGTGG CATTCACTCC 780
CTGTATGGAG ACCCAAAAGA GAACCAACGC TTGCCAAATC CTGACAAATC AGAACAGCT 840
CTCTGTGACC CCAATTGAG TTTTGATGCT GTCACACCG TGGGAAATAA GATCTTTTC 900
TTCAAAGACA GGTCTCTCTG GCTGAAGGTT TCTGAGAGAC CAAAGACCA: TGTTAATTTA 960
ATTTCTCTCT TATGGCCAAC CTGCGCATCT GGCATTGAAG CTGCTTATGA AATTGAAGCC 1020
AGAAATCAAG TTTTCTTTT TAAAGATGAC AAATCTGGT TAATTAGCAA TTTAAGACCA 1080
GAGCCAAATT ATCCCAAGAG CATACATTCT TTTGGTTTC TAACTTTGT GAAAAAATT 1140
GATGCAGCTG TTTTAAACCC ACGTTTAT AGGACCTACT TCTTTGAGA TAACCAAGTAT 1200
TGGAGGTATG ATGAAAGGAG ACAGATGATG GACCTGGTGT ATCCCAACT GATTACCAAG 1260
AACTTCAAAG GAATCGGGCC TAAATTGAT GCAGTCTTCT ACTCTAAAA CAAATACTAC 1320
TATTTCTTCC AAGGATCTAA CCAATTGAA TATGACTTCC TACTCCAAG TATCACCAAA 1380
ACACTGAAA GCAATAGCTG GTTTGGTGT TGA

Seq ID NO: 525 Protein sequence
Protein Accession #: P39900

85

1 11 21 31 41 51
MKPLLILLQ ATASGALPLN SSTSLKKNV LFGERYLEK YGLEINKLPV TKMKYSGNLM 60
KEKIQEMQHP LGLKVTGQLD TSTLEMMHAP RGVDPDVHVF REMPGGPVWR KHYITYRINN 120

YTPDMNREDV DYAIRKAFQV WSNVTPLKPS KINTGMADIL VVFARGAGD FHAFDGGKGI 180
 LAHAFGPQSC IGGDAHDEED EFWTHSGGT NLFLTAVHEI GHSGLGHSS DPKAVMFPTY 240
 KYVDINTFRL SADDIRGIQS LYGDPKENQR LPNPDNSEPA LCDPNLSFDA VTTVGNKIPF 300
 FKDRFFWLKV SERPKTSVNL ISSLWPTLPS GIEAAEIEA RNQVFLPKDD KYWLISNLRP 360
 EPNYPKSIHS FGFPNFKVKI DAAVFNPRFY RTYFFVDNQY WRDERRQMM DPGYPKLITK 420
 NFGIGPKID AVFYSKNKY YFFQGSNQFE YDFLLQRITK TLKSNSWFGC

Seq ID NO: 526 DNA sequence

Nucleic Acid Accession #: NM_024423.1

Coding sequence: 64..2590

1 11 21 31 41 51
 15 GGCAGGCTCT GCTCTCGGCA CCTCCCGGCG GCGCGCGTTC TCCTGGCCCT GCCCGGCATC 60
 CCGATGGCCG CCGCTGGGCC CCGGCGCTCC GTGCGCGGAG CCGTCTGCCT GCATCTGCTG 120
 CTGACCCCTCG TGAFTCTCAG TOGTGATGGT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180
 CCTTCTAAAC TAGAGGCAGA CAAAATAATT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240
 TCTGCAGACC TCATCCGGTC AAGTGATCCT GATTTCAGAG TICTAAATGA TGGGTCAGTG 300
 20 TACACAGCCA GGGCTGTGTC GCTGTCTGAT AAGAAAAGAT CATTACCAT ATGGCTTTCT 360
 GACAAAAGGA AACAGACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420
 TCGAAGACAA GACACACTAG AGAAACTGTT CTCAGGCGTG CCAAGAGGAG ATGGGCACCT 480
 ATTCTTGCT CTATGCAAGA GAATTCCTTG GGCCTTTCC CATGTTTCT TCAACAAGTT 540
 GAATCTGATG CAGCACAGAA CTATCTGTC TTCTACTCAA TAAGTGAAG TGGAGTTGAT 600
 AAAGAACTT TAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTATT TTGCACTCGG 660
 25 CCTGTGGATC GTGAAGATA TGTATTTTT GATTGTATG CTATGCGTC AACTGCAGAT 720
 GGATATTACG CAGATCTGCC CCTCCACTA CCCATCAGGG TAGAGGATGA AATGACAAC 780
 CACCCTGTTT TCACAGAAGC AATTATATAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840
 ACTCAGTGGG GGGTGGTTTG TGCCACAGAC AGAGATGAAC CGACACAAT GCATACGCGC 900
 CTGAATATACA GCATTITGCA GCAGACACCA AGGTCACCTG GGCTCTTTTC TGTGCATCCC 960
 30 AGCAGCGGCG TAATCACCAC AGTCTCTCAT TATTGGACA GAGAGGTGTG AGACAAGTAC 1020
 TCATTGATAA TGAAGATACA AGACATGGAT GGCCAGTTTT TTGATTGAT AGGCACATCA 1080
 ACTTGATATCA TAACAGTAAC AGATTCAAAT GATAATGCAC CCACCTTCAG ACACAAATGCT 1140
 TATGAAGCAT TTGTAGAGGA AAATGCATTG AATGTGGAAT TCTTACGAAT ACCTATAGAA 1200
 GATAAGGATT TAAATTAACAC TGCCAAATTGG AGAGTCAATT TTACCAITTT AAAGGGAAT 1260
 35 GAAATGGAG ATTTCAAAT CAGCACAGAC AAAGAACTA ATGAAGGTGT TCTTTCTGTT 1320
 GTAAAGCCAC GGAATTATGA AGAAAACCGT CAAGTGAACC TGGAAATGG AGTAAACAAT 1380
 GAAGCGCCAT TTGCTAGAGA TATCCACAGA GTGACAGCCT TGAACAGAGC CTGGTTTACA 1440
 GTTCATGTGA GGGATCTGGA TGAGGGGCGCT GAATGCACCT CTGCAGCCCA ATATGTGCGG 1500
 ATTAAGAAAT ACTTAGCAGT GGGGTCAAAG ATCAACGGCT ATAAGGCATA TGACCCCGAA 1560
 40 AATAGAAATG GCAATGGTTT AAGGTACAAA AAATTCATG ATCCTAAAGG TTGGATCACC 1620
 ATTGATGAAA TTTCAGGATC AATCATAACT TCCAAATCC TGGATAGGGA GGTGAAACT 1680
 CCCAAAATG AGTTGTATAA TATTACAGTC CTGGCAATAG ACAAGATGA TAGATCATGT 1740
 ACTGGAACAC TTGCTGTGAA CATTGAAGAT GTAATGATA ATCCACCAGA AATACTTCAA 1800
 GAATATGTAG TCAATTGCAA ACCAAAATG GGGTATACCG ACATTTTAGC TGTGTATCCT 1860
 45 GATGAACCTG TCCATGGAGC TCCATTTTAT TTCAGTTTGC CCAATACTTC TCCAGAAATC 1920
 AGTAGACTGT GGAGCCTCAC CAAAGTTAAT GATACAGCTG CCGCTCTTTC ATATCAGAAA 1980
 AATGCTGGAT TTCAAGATA TACCATTCTT ATTACTGTAA AAGACAGGGC CGGCCAAGCT 2040
 GCAACAAAT TATTGAGAGT TAATCTGTGT GAATGTACTC ATCCAACTCA GTGTGCTGG 2100
 ACTTCAAGGA GTACAGGAGT AATCTTGGTA AAATGGGCAA TCCTTGCAAT ATTACTGGGT 2160
 50 ATAGCAGTCG TCTTTCTGTT ATTGCTAACT TTAGTATGTG GAGTTTTTGG TGCAACTAAA 2220
 GGGAAACGTT TTCTGAAAG TTTAGCACAG CAAAACCTTA TTATATCAAA CACAGAAGCA 2280
 CCTGGAGACG ATAGAGTGTG CTCTGCCAAT GGAATTTATG CCAAACTAC CAACAACCTC 2340
 AGCCAAGGTT TTTGTGGTAC TATGGGATCA GGAATGAAAA ATGGAGGCA GGAACCAATT 2400
 GAAATGATGA AAGGAGGAAA CCAGACCTTG GAATCCTGCC GGGGGGCTGG GCATCATCAT 2460
 55 ACCCTGGACT CCTGCAGGGG AGGACACACG GAGGTGGACA ACTGCAGATA CACTTACTCG 2520
 GAGTGGCACA GTTTTACTCA ACCCGCTCTC GGTGAAGAAT CCATTAGAGG ACACACTGGT 2580
 TAAAAATTA ACATAAAGA AATTGCATCG ATGTAATCAG AATGAAGACC GCATGCCATC 2640
 CCAAGATTAT GTCCCTCACT ATAATATGA GGAAGAGGA TCTCCAGCTG GTTCTGTGGG 2700
 CTGCTGCAGT GAAAGCAGG AAGAAGATGG CCTTGACTTT TTAATAAATT TGAACCCAA 2760
 60 ATTTATTACA TTAGCAGAAG CATGCACAAA GAGATAATGT CACAGTGCTA CAATTAGGTC 2820
 TTTGTGACAC ATTCTGGAGG TTTCCAAAA TAATATTGTA AAGTTCAATT TCAACATGTA 2880
 TGTATATGAT GATTTTTTTC TCAATTTTGA ATTATGCTAC TCAACCAATT ATATTTTTAA 2940
 AGCCAGTTGT TGCTTATCTT TTCCAAAAAG TGAATAATGT TAARAACAG AACTGGTAAA 3000
 TCTCAAACTC CAGCACTGGA ATTAAGGTCT CTAAGGCATC TGCTCTTTTT TTTTTTTACG 3060
 65 GATATTTTAG TAAATAATAT GCTGGATAAA TATTAGTCCA ACATAGCTA AGTTATGCTA 3120
 ATATCACATT ATTATGATT CACTTAAAGT GATAGTTTAA AAAATAAACA AGAAATATTG 3180
 AGTATCACTA TGTGAAGAAA GTTTTGGAAA AGAAACATG AAGACTGAAT TAAATTAATA 3240
 ATGTTGCAGC TCATAAGAA TTGGGACTCA CCCCTACTGC ACTACCAAT TCAATTGACT 3300
 TTGGAGGCAA AATGTGTTGA AGTGCCCTAT GAAGTAGCAA TTTCTATAG GAATATAGTT 3360
 70 GGAATAAAT GTGTGTGTGT ATATTATTAT TAATCAATGC AATATTTAAA ATGAAATGAG 3420
 AACAAAGAGG AAAATGGTAA AAACCTGAAA TGAGGCTGGG GTATAGTTTG TCCTACATA 3480
 GAAAAAGAG AGAGCTTCCT AGGCCTGGSC TCTTAAATGC TGCAATTATA CTGAGTCTAT 3540
 GAGGAAATAG TTCTGTGCTA ATTTGTGTAA TTTGTTTAAA ATTGTAAATA AATTAACTT 3600
 75 TCTGTGTTTC TGTGGGAGG AAATAGGGAA TCCAATGGAA CAGTAGCTTT GCTTTGCACT 3660
 CTGTTTCAAG ATTTCTGCAT CCACAAGTTA GTAGCAAACT GGGGAATACT CGCTGCAGCT 3720
 GGGTTCCTCT GCTTTTGGT AGCAAGGGTC CAGAGATGAG GTGTTTTTTT CGGGGAGCTA 3780
 ATAACAAAA CATTAAAAA CTTACCTTTA CTGAAGTTAA ATCCTCTATT GCTGTTTCTA 3840
 TTCTCTCTTA TAGTGACCAA CATCTTTTAA ATTTAGATCC AAATAACCAT GTCTCTTAG 3900
 80 AGTTTAGAGG CTAGAGGGAG CTGAGGGGAG GATCTTACTG AAAGCACCCT GGGGAGATTG 3960
 ATTTGCTCTA AACCTAAGCC CCACAACTT GACACCTGAT CAGGTCTGGG AGCTACAAAA 4020
 TTCTCATTTT CTCTCACTG CCCTTCTTCT GAGTGGCAAT GGCCCTGAAT AAGGAAAGCC 4080
 AGGCCTTGTG GSCCCCTTCT TTTGGGCTTT CTGCTAAAGC AACACCTCCA GCAGAGATT 4140
 CCTTAAGTGA CTCACAGTTT TCCACCATCC TTCACGCTGA ATTAATTTTT AATCAGTTT 4200
 85 CTTTCTCCAG AGAAATTTTA AAATAATAGA AGAAATAGAA ATTTTGAATG TATAAAGAA 4260
 AAAGATCAAG TTGTCATTTT AGAACAGAGG GAACCTTGGG AGAAAGCAGC CCAAGTAGGT 4320
 TATTTGTACA GTACAGGGCC AACAGGAAGA TGACGGCTTT CAAGGGCAAG GAGAGGCCAC 4380
 AAGGAATATG GGTGGGAGTA AAAGCAACAT CGTCTGCTTC ATACTTTTTT CTAGGCTTGG 4440

5
10
15
20
25
30
35
40
45

CACATGCTTT TCCCTTCTCA GGCCAATGGC AACTGCCATT TGAGTCCGGT GAGGATCAG 4500
 CCAACCTCTT CTCTATGGCT CACCTTATTT GGAGTGAGAA ATCAAGGAGA CAGAGCTGAC 4560
 TGCAATGATG GTCTGAAGGC ATTTGAGGA TGAGCCTGAA CTGGTTGTGC AGAACAAACA 4620
 AGGCATTCAT GGGAAATGTT GTATTCTCTC TGCAGCCCTC CTCTCTGGCA CTAAGAAAGT 4680
 CTATGAATTA AATGCCATAT TAAATCTCTG ATTTATTCCT ACATTTCTCT TTTCTTAATT 4740
 TGACCTTAAA ATCTATGTGT TTTAGACTTA GACTTTTATAT TGCCCCCCCC CCCTTTTTTT 4800
 TTGAGACGGA GTCTCGCTCT GAOCACAGG CTGGAGTGCA GTGGCTCCGA TCTCTGCTCA 4860
 CTGAAGAGTC GGCCTCCCGG GTTCATGCCA TTCTCTGGCC TCAGCCTCCT GAGTAGCTGG 4920
 GACTAGAGGC GCCCAACCAC ACGCCCGGCT AATTTTTTGT ATTTTAAATA GAGACGGGGT 4980
 TTCCTGTGT TAGCCAGGAT GGTCTCGATC TCCTGACCTC GTGATCCGCC TGCCCTCGGC 5040
 TCCCAAAGTG CTGGGATTAC AGGCATGACC CACCCTCCCC GGCCTTGTTC TCCGTTTAAA 5100
 GTCGCTCTCT TTTAATGTAA TCATTTGAA CATGTGTGAA AGTTGATCAT ACGAATTGGA 5160
 TCAATCTTGA AATACTCAAC CAAAAGACAG TCGAGAAGCC AGGGGGAGAA AGAACTCAGG 5220
 GCACAAAATA TTGCTCTGAG AATGGAATTC TCTGTAAGCC TAGTTGCTGA AATTCTCTGC 5280
 TGTAAACAGA AGCCAGTTTT ATCTAACGGC TACTGAAACA CCCACTGTGT TTTGCTCACT 5340
 CCCACTCACC GATCAAAACC TGCTACCTCC CCAAGACTTT ACTAGTCCCG ATAAACTTTC 5400
 TCAAGAGGCA ACCAGTATCA CTTCCTGTT TATAAAACCT CTAACCATCT CTTTGTCTCT 5460
 TGAACATGCT GAACAACCAT TGGTCTGCAT GTATGCCCGA ATTTGTAATT CTTTCTCTC 5520
 AAATGAAAT TTAATTTTAG GGATTCATTT CTATATTTTC ACATATGTAG TATTATTATT 5580
 TCCITATATG TGTAAGGTGA AATTTATGGT ATTTGAGTGT GCAAGAAAAT ATATTTTAA 5640
 AGCTTTTCAAT TTTCCCCCAG TGAATGATTT AGAATTTTTT ATGTAATATAT ACAGAATGTT 5700
 TTTTCTTACT TTTATAAGGA AGCAGCTGTC TAAATGTCAG TGGGTTTGT TTTGCAATGT 5760
 TTTAAACAGA TTTTGTAGTAT TGCTATTAAA AGAAGTTACT TTGCTTTTAA AGAACTTGG 5820
 CTGCTTAAAA TAAGCAAAAA TTGGATGCAT AAAGTAATAT TTACAGATGT GGGGAGATGT 5880
 AATAAAACAA TATTAACCTG GCTGCTTAAA ATAAGCAAAA ATTGGATGCA TAAAGTAATA 5940
 TTTACAGATG TGGGGAGATG TAATAAAACA ATATTAACCT GGTTCCTGT TTTGCTGTA 6000
 TTTAGAGATT AAATAATTTCT AAGATGATCA CTTTGCAAAA TTATGCTTAT GCGTGGCATG 6060
 GAAATAGAAA TACTCAATTA TGTCTTGTG GTATTAATGG GGAATATTTT GGACAATGTT 6120
 TCATTATCAA ATTGTCGACA TCATTAAATAT ATATTGTAAT GTTGGGAAGA GATCACTATT 6180
 TTGAAGCACA GCTTTACAGA TGAGTATCTA TGATACATAT GTATAATAAA TTTTGTATCG 6240
 GTATTAAAAG TATTAGAAGG TGGTTATAAT TGCAGAGTAT TCCATGAATA GTCACTGAC 6300
 ACAGGGGTTT TACTTTGAGG ACCAGTGTAG TCAAGGGAAA ACATGAGTTA AAAAGAAAG 6360
 CAGGCAATAT TGCAGTCTTG ATTCTGCCAC TTACAGGATA GATAATGCCT GAACCTTAA 6420
 GACAAGATGA TCCAACATA AAGGTGCTCT GTGCTTACA GTGAATCTTT TCCCATGCA 6480
 GGAGTGTGCT CCCCTACAAA CGTTAAGACT GATCATTTCA AAAATCTATT AGCTATATCA 6540
 AAAGCCTTAC ATTTTAATAT AGGTTGAACC AAAATTTCAA TTCCAGTAAC TTCTATTGTA 6600
 ACCATTATTT TTGTGTATGT CTTCAAGAAAT GTTCATTGGA TTTTGTGTTG TAATAGTAAA 6660
 ATACCGGATA CATTTCAGT GTCCCTCAGT ATTGATTGG TTGAATATTG GGTCAATAG 6720
 GTTGAGAACG ATGGACACTA GAGCCAGAAAT GCTTGGATAT GAATCCTGGA TCTGTCACT 6780
 ACTTCTGTGT GACCTTTGAA AGGCTACTTA TTTCTCTCT TAGCTTTCTC ATTAAATCA 6840
 ATGAACAATG CCAGCCTCAT GGGGTTGTTG AATGATTAAA TTAGTTAATA TACCTAAAGT 6900
 ACATAGAACA CTGCCTGCAC ATAGTAAAG AATTATAAGT GTGAGGTAGT TGGTAAATTT 6960
 ATGTAGTTGG ATATACTACC GAACAATATC TAATCTCTTT TTAGGGAAAT AAAGTTTGTG 7020
 CATATATATA ATCCGGAAC ATG

Seq ID NO: 527 Protein sequence
 Protein Accession #: NP_077741.1

50
55
60
65

1 11 21 31 41 51
 MAAAGPRRSV RGAVCLHLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECFRS 60
 ADLIRSDPD PRVLNDGSVY TARAVALS DK KRSFTIWLSD KRKQTQKEVT VLEHQKKVS 120
 KTRHRETVEL RRAKRRWAPI PCSMQENSLG PPFLFLQQVE SDAAGNYTVF YSISGRGVDK 180
 EPLNLFYIER DTGNLFCTRP VDREYDVFD LIAAYASTADG YSADLPLPLP IRVEDENDNH 240
 PVFTEAIYNF EVLESSRPGT TVGVVCAADR DEPDTHMTRL KYSILQOTPR SPGLFSVHPS 300
 TGVITTVSHY LDREVVDKYS LIMKVQMDMG QFFGLIGTST CIITVDSND NAPTFRQNAV 360
 EAFVEENAFN VEILRIPIED KDILINTANWR VNFITLKNGE NGHFKISTDK ETNEGVLVSV 420
 KPLNYEENRN VNLLEIGVNE APFARDIPRV TALNRLVTV HVRDLDEGPE CTPAAQYVRI 480
 KENLAUGSKI NGYKAYDPEN RGNGLRYKK LHPDKGWIT DEISGSIITS KILDREVETP 540
 KNELNYITVL AIDKDRSCT GLAVNIEDV NDNPPHILQE YVVICPKMG YTDILAVDPD 600
 EPVHGAPFPY SLPNTSPBIS RLWSLTKVND TAARLSYQKN AGFQYETIPI TVKDRAGQAA 660
 TKLLRVNLCE CTHPTQCRAT SRSTGVILGK WAILAILLGI ALLFSVLLTL VCGVFGATKG 720
 KRFPEDLAQQ NLIISNTEAP GDDRVCASNG FMTQTNNS QFGCGTMGSG MKNGGQETIE 780
 MMRGNQITL SCRGAGHHT LDSCRGGHTE VDNCRITYSE WBSFTQPRLG EESIRGHTG

Seq ID NO: 528 DNA sequence
 Nucleic Acid Accession #: NM_001941.2
 Coding sequence: 64..2754

70
75
80
85

1 11 21 31 41 51
 GGCAGGCTCT GCTCTCGGCA CCCTCCCGGC GCCCGCGTTC TCCTGGCCCT GCCCGGCATC 60
 CCGATGGCGG CCGCTGGGCC CCGGCGCTCC GTGCGCGGAG CCGCTGCTCT GCATCTGCTG 120
 CTGACCCCTG TGATCTTCAG TCGTGATGGT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180
 CCTTCTAAAC TAGAGGCAGA CAAAATAATT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240
 TCTGAGACCC TCATCCGCTC AAGTGATCCT GATTTCAGAG TTCTAAATGA TGGTCTCAGT 300
 TACACAGCCA GGGCTGTGTC GCTGTCTGAT AAGAAAAGAT CATTTACCAT ATGGCTTTCT 360
 GACAAAAGGA AACAGACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420
 TCGAAGACAA GACACACTAG AGAACTGTT CTCAGGCGTG CCAAGAGGAG ATGGGCACCT 480
 ATTCCTTGCT CTATGCAAGA GAATTCCTTG GGCCTTTTCC CATTGTTTCT TCAACAAGTT 540
 GAATCTGATG CAGCACAGAA CTATACTGTC TTCTACTCAA TAAGTGGACG TGGAGTTGAT 600
 AAAGAACCTT GAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTATT TTCACTCGG 660
 CCTGTGGATC TGTGAAGAATA TGATGTTTTT GATTTGATTG CTTATGCGTC AACTGCAGAT 720
 GGATATTTCAG CAGATCTGCC CCTCCCACTA CCCATCAGGG TAGAGGATGA AAATGACAAC 780
 CACCCCTGTT TCACAGAAGC AATTATATAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840
 ACTACAGTGG GGTGGTTTGT TGCCACAGAC AGAGATGAAC CGGACACAAT GCATACGCGC 900
 CTGAAATACA GCATTTTGCA GCAGACACCA AGGTCACCTG GGCCTCTTTC TGTGCATCCC 960

	AGCACAGGCG	TAATCACCAC	AGTCTCTCAT	TATTTGGACA	GAGAGGTTGT	AGACAAGTAC	1020
	TCATTGATAA	TGAAAGTACA	AGACATGGAT	GGCCAGTTT	TTGGATTGAT	AGGCACATCA	1080
	ACTTGATATCA	TAACAGTAAAC	AGATTCAAAT	GATAATGCAC	CCACTTTTACG	ACAAAATGCT	1140
	TATGAAGCAT	TTGTAGAGGA	AAATGCATT	AATGTGGAAA	TCTTACGAAT	ACCTATAGAA	1200
5	GATAAGGATT	TAATTAACAC	TGCCAATTGG	AGAGTCAATT	TTACCATTTT	AAAGGGAAAT	1260
	GAAAATGGAC	ATTTCAAAT	CAGCAGACAC	AAAGAACTA	ATGAAGGTTG	TCTTTCTGTT	1320
	GTAAAGCCAC	TGAATTATGA	AGAAACCGT	CAAGTGAACC	TGGAAATTGG	AGTAAACAAAT	1380
	GAAGCGCCAT	TTGCTAGAGA	TATTCACAGA	GTGACAGCCT	TGAACAGAGC	CTTGTTTACA	1440
	GTTCATGTGA	GGGATCTGGA	TGAGGGGCT	GAATGCATC	CTGCAGCCCA	ATATGTGGGG	1500
10	ATTAAGAAA	ACTTAGCAGT	GGGTCAAAG	ATCAACGGCT	ATAAGGCATA	TGACCCCGAA	1560
	AATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTGCATG	ATCCTAAAGG	TTGGATCACC	1620
	ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GGTTGAAACT	1680
	CCCAAAATG	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
	ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAATGATA	ATCCACCAGA	AATACTTCAA	1800
15	GAATATGTAG	TCATTTGCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
	GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTCAAGTTTC	CCAATACITC	TCCAGAAATC	1920
	AGTAGACTGT	GGAGCCTCAC	CAAAAGTTAAT	GATACAGCTG	CCCGCTTTTC	ATATCAGAAA	1980
	AATGCTGGAT	TTCAAGATA	TACCATTCTT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
	GCAACAAAAT	TATTCAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACCTCA	GTGCTGTGGC	2100
20	ACTTCAAGGA	GTACAGGAGT	AATACTTGGG	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
	ATAGCATGTC	TCTTTTCTGT	ATTGCTAACT	TTAGTATGTG	GAGTTTITGG	TGCAACTAAA	2220
	GGGAAAACGT	TTCTGTAAGA	TTTAGCACAG	CAAAACTTAA	TTATATCAAA	CACAGAAAGCA	2280
	CTCGGAGACG	ATAGAGTGTG	CTCTGCCAAT	GGATTATGTA	CCCAAACTAC	CAACAACCTCT	2340
	AGCCAAAGTT	TTGCTGTGAA	TATGGGATCA	GGAATGAAAA	ATGGAGGGCA	GGAACCAATT	2400
25	GAAATGATGA	AAGGAGGAAA	CCAGACCTTG	GAATCTGCCC	GGGGGGCTGG	GCATCATCAT	2460
	ACCCCTGGACT	CTCGCAGGGG	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACITACTCG	2520
	GAGTGGGACA	GTTTTACTCA	ACCCCGTCTC	GGTGAATAAT	TGCATCGATG	TAATCAGAAT	2580
	GAAGACCGCA	TGCCATCCCA	AGATTATGTC	CTCACTTATA	ACTATGAGGG	AAGAGGATCT	2640
	CCAGCTGGTT	CTGTGGGCTG	CTGCAGTGAA	AAGCAGGAAG	AAGATGGCCT	TGACTTTTTA	2700
30	AATAATTGGG	AACCCAAATT	TATTACATTA	GCAGAAGCAT	GCACAAAGAG	ATAATGTCAC	2760
	AGTGCTACAA	TTAGGTCCTT	GTGAGACATT	CTGGAGGTTT	CCAAAAATAA	TATTGTAAAG	2820
	TTCAATTCCA	ACATGTATGT	ATATGATGAT	TTTTTTCTCA	ATTTTGAATT	ATGCTACTCA	2880
	CCAATTATTA	TTTTTAAAGC	CAGTTGTGTC	TTATCTTTTC	CAAAAAGTGA	AAAAATGTTA	2940
	AACAGACAAC	TGGTAAATCT	CAAACTCCAG	CACCTGGAAT	AAGGTCTCTA	AAGCATCTGC	3000
35	TCTTTTTTTT	TTTTACGGAT	ATTTTAGTAA	TAAATATGCT	GGATAAATAT	TAGTCCAACA	3060
	ATAGCTAAGT	TATGCTAATA	TCACATTATT	ATGTATTAC	TTTAAGTGAT	AGTTTAAAAA	3120
	ATAAACAGAA	AATATTGAAT	ATCACTATGT	GAAGAAAGTT	TTGGAAAAGA	AACAATGAAG	3180
	ACTGAATTAA	ATTAATAATG	TTGCAGCTCA	TAAAGAAATG	GGACTCACCC	CTACTGCAC	3240
	ACCAAAATTC	TTGACTTTTG	GAGGCAAAAT	GTGTTGAAGT	GCCCTATGAA	GTAGCAATTT	3300
40	TCTATAGGAA	TATAGTTGGA	AATAAATGTG	TGTGTGTATA	TTATTATTA	TCAATGCAAT	3360
	ATTTAAAAAT	AAATAGAGAC	AAAGAGGAAA	ATGTTAAAAA	CTTGAAATGA	GGCTGGGGTA	3420
	TAGTTTGTCC	TACAATAGAA	AAAAGAGAGA	GCTTCTTAGG	CCTGGGCTCT	TAAATGCTGC	3480
	ATTATPACTG	AGTCTATAG	GAAATAGTTC	CTGTCCAATT	TGTGTAATTT	GTTTAAAAAT	3540
	GTAATAAAT	TAAACTTTTC	TGGTTTCTGT	GGGAAGGAAA	TAGGGAATCC	AATGGAAACG	3600
45	TAGCTTTGCT	TTGCAGTCTG	TTTCAAGATT	TCTGCATCCA	CAAGTTAGTA	GCAAACTGGG	3660
	GAATACTCGG	TGCAGCTGGG	GTTCCTGCT	TTTGGTAGC	AAGGGTCCAG	AGATGAGGTG	3720
	TTTTTTTCGG	GGAGCTAATA	ACAAAAACAT	TTTAAAACTT	ACCTTTACTG	AAGTTAAATC	3780
	CTCTATGTCT	GTTCCTATTC	TCTCTTAG	TGACCAACAT	CTTTTAAAT	TAGATCCAAA	3840
	TAACCATGTC	CTCCTAGAGT	TTAGAGGCTA	GAGGGAGCTG	AGGGGAGGAT	CTTACTGAAA	3900
50	GCACCCTGGG	GAGATTGATT	GTCTTAAAC	CTAAGCCCCA	CAAACTTGAC	ACCTGATCAG	3960
	GTCTGGGAGC	TACAAAATTT	CATTTTCTC	CTCACTGCC	TTCTTCTGAG	TGGCATTTGG	4020
	CTGAATCAAG	GAAAGCCAGG	CCCTGTGGGC	CCCTTCTTT	CGGCTTCTG	CTAAAGCAAC	4080
	ACCTCCAGCA	GAGATTCCCT	TAAGTGACTC	CAGGTTTCC	ACCATCCTTC	AGCGTGAATT	4140
	AAITTTTAAAT	CGTTTGTCTT	TCTCCAGAGA	AATTTTAAAA	TAATAGAAGA	AATAGAAAT	4200
55	TTGAATGTAT	AAAAGAAAAA	GATCAAGTTG	TCATTTTAGA	ACAGAGGGAA	CTTTGGGAGA	4260
	AAGCAGCCCA	AGTAGGTTAT	TTGTACAGTC	AGAGGGCAAC	AGGAAGATGC	AGGCCTTCAA	4320
	GGGCAAGGAG	AGGCCACAAG	GAATATGGGT	GGGAGTAAAA	GCAACATCGT	CTGCTTCATA	4380
	CTTTTCTCTA	GGCTTGGCAC	TGCTTTTCC	TTTCTCAGGC	CAATGGCAAC	TGCCATTGGA	4440
	GTCCGGTGAG	GGATCAGCCA	ACCTCTTCTC	TATGGCTCAC	CTTATTTTGA	GTGAGAAATC	4500
60	AAGGAGACAG	AGCTGACTGC	ATGATGAGTC	TGAAGGCATT	TGCAGGATGA	GCCTGAACTG	4560
	GTGTGTCAGA	ACAAACAAGG	CATTCATGGG	AATGTGTGTA	TTCTTCTG	AGCCCTCTCT	4620
	CTGGGCACTA	AGAAGTCTTA	TGAATTAAAT	GCCTATCTAA	AATCTGTGAT	TATCTCTACA	4680
	TTTTCTGTTT	TCTAATTTGA	CCCTAAATTC	TATGTGTTTT	AGACTTAGAC	TTTTTATTGC	4740
	CCCCCCCCCT	TTTTTTTTTT	AGACGGAGTC	TGCTCTGAC	GCACAGGCTG	GAGTGCAGTG	4800
65	GCTCCGATCT	CTGCTCACTG	AAAGCTCCGC	CTCCCGGGTT	CATGCCATT	TCCTGCCTCA	4860
	GCCTCCTGAG	TAGCTGGGAC	TACAGGCGCC	CACCACCACG	CCCGGCTAAT	TTTTTGTATT	4920
	TTTAATAGAG	ACGGGGTTTC	ACTGTGTTAG	CCAGGATGGT	CTCGATCTCC	TGACCTCGTG	4980
	ATCCCGCTGC	CTGGGCTCC	CAAGTGCTG	GGATTACAGG	CATGACCCAC	CGCTCCCGGC	5040
	CTGTGTTTCC	GTTTAAAGTC	GTCTTCTTTT	AATGTAATCA	TTTGAACAT	GTGTGAAAGT	5100
70	TGATCATACG	AATTTGGATCA	ATCTTGAAT	ACTCAACCAA	AAGACAGTCG	AGAAGCCAGG	5160
	GGGAGAAAGA	ACTCAGGGCA	CAAAATATTG	GTCTGAGAAT	GGAATTTCT	GTAAGCCTAG	5220
	TTGCTGAAAT	TTCTGCTGT	AACCAGAAAG	CAGTTTTATC	TAACGGCTAC	TGAACACCCC	5280
	ACTGTGTTTT	GCTCACTCCC	TCACTCACCG	ATCAAAACCT	GCTACCTCCC	CAAGACTTTA	5340
	CTAGTGCCGA	TAACTTTTCT	CAAGAGCAAA	CCAGTATCAC	TTCCCTGTTT	ATAAAACCTC	5400
75	TAACCATCTC	TTTGTCTTTT	GAACATGCTG	AAAACCACTT	GGTCTGCATG	TATGCCCGAA	5460
	TTTGTAAATC	TTTTCTCTCA	AATGAAAAAT	TAATTTTAGG	GATTCAATTC	TATATTTTCA	5520
	CATATGATGT	ATTATTATGT	CTTATATATG	GTAAAGGTGAA	ATTTATGGTA	TTTGAGTGTG	5580
	CAAGAAAAAT	TATTTTAAAA	GCITTCATTT	TTCCCCAGT	GAATGATTTA	GAATTTTTTA	5640
	TGTAATATATA	CAGAATGTTT	TTTCTTACTT	TTATAAGGAA	GCAGCTGTCT	AAAATGCAGT	5700
80	GGGGTTTGGT	TTGCAATGTT	TTTAAACAGAG	TTTTAGTATT	GCTATTAATA	GAAGTTACTT	5760
	TGCTTTTAAA	GAAACTTGGC	TGCTTAAAAAT	AAGCAAAAAT	TGGATGCATA	AAGTAATATT	5820
	TACAGATGTG	GGGAGATGTA	ATAAAACAAAT	ATTAACCTGG	TTTCTGTTT	TTGCTGATT	5880
	TAGAGATTAA	ATAATTCTAA	GATGATCACT	TTGCAAAATT	ATGCTTATGG	CTGGCATGGA	5940
	AATAGAAAAA	CTCAATTATG	TCTTTTGTGT	ATTAATGGGG	AATATTTTGG	ACAATGTTTC	6000
85	ATTATCAAAAT	TGTCGACATC	ATTAATATAT	ATTGTAATGT	TGGGAAGAGA	TCACTATTTT	6060
	GAAGCACAGC	TTTACAGATG	AGTATCTATG	ATACATATGT	ATAATAAATT	TTGATCGGGT	6120
	ATTAAGAGTA	TTAGAAGGTG	GTTATAATTG	CAGAGTATTC	CATGAATAGT	ACACTGACAC	6180

WO 02/086443

5 AGGGGTTTAA CTTTGGAGAC CAGTGTAGTC AAGGGAAAAC ATGAGTTAAA AAGAAAAGCA 6240
 GGCAATATTG CAGTCTTGAT TCTGCCACTT ACAGGATAGA TAATGCCTGA ACTTTAATGA 6300
 CAAGATGATC CAACCATAAA GGTGCTCTGT GCTTCACAGT GAATCTTTTC CCCATGCAGG 6360
 AGTGTGCTCC CCTACAAACG TTAAGACTGA TCATTTCAA AATCTATTAG CTATATCAAA 6420
 AGCCTTACAT TTTAATATAG GTTGAACCAA AATTCAATT CCAGTAACCT CTATTGTAAC 6480
 CATTATTTT GTGTATGTCT TCAAGAATGT TCATTGGATT TTTGTTTGA ATAGTAAAT 6540
 ACCGGATACA TTCACTGTGT CCTTCAGTAT TGATTGGTT GAATATTGGG TCATAATGGT 6600
 TGAGAAGCAT GCACACTAGA GCCAGAATGC TTGGATATGA ATCCTGGATC TGTCACTTAC 6660
 10 TTCTGTGTGA CCTTTGAAGG GCTACTTATT TCCTCTCTTA GCTTCTCAT TAAATCAAT 6720
 GAACAATGCC AGCCTCATGG GGTGTGTGAA TGATTAAAT AGTTAATATA CCTAAAGTAC 6780
 ATAGAACACT GCCTGCACAT AGTAAAAGAA TTATAAGTGT GAGGTAGTTG GTAAAATTAT 6840
 GTAGTTGGAT ATACTACCGA ACAAATATCT ATCTCTTTT AGGGAATAA AGTTTGTGCA 6900
 TATATATAAT CCGAAACAT G

15 Seq ID NO: 529 Protein sequence
 Protein Accession #: NP_001932.1

1 11 21 31 41 51
 20 MAAAGPRRSV RGAVCLHLL TLVIFSRDGE ACKKVLNVLP SKLEADKIIG RVNLEECFRS 60
 ADLIRSSDPD FRVLNDGSVY TARAVALSDK KRSPTIWLSD KRRQTQKEVT VLEHQKKVS 120
 KTRHTRETVL RRAKRRWAPI PCSMQENSLG PFPLFLQQVE SDAQNYTVF YSISGRGVDK 180
 EPLNLFYIER DGNLFCFTRP VDREYDVFD LIAYASTADG YSADLPLPLP IRVEDENDNH 240
 25 PVFTEAIYNP GVELESSRPGT TVGVVCATDR DEPTMTHTRL KYSILQOTPR SPGLFSVHPS 300
 TGVITTVSHY LDREVVDKYS LIMKVQDMDG QPFGLIGTST CIITVTDSDND NAPTFRQNAV 360
 EAFVEENAFN VEILRIPIED KDLINTANWR VNFTILKNGE NGHFKISTDK ETNEGVLVSV 420
 KPLNYENRQ VNLSEIGVNE APPARDIPRV TALNRALVTV HVRDLDEGPE CTPAAQYVRI 480
 KENLAVGSKI NGYKAYDPEN RENGGLRYKK LHPDKGWITI DEISGSIITS KILDREVETP 540
 KNELYNITVL AIDKDRSCT GTLAVNIEDV NDNPEILQE YVICKPKMG YTDILAVDPD 600
 30 EPVHGAPFFP CTHPNTSPSIS RLWSLTKVND TAARLSYQKN AGFQYETIPI TVKDRAGQAA 660
 TKLLRVNLCE SLHPTQCRAT SRSTGVILGR WAILAILLGI ALLPSVLTL VCGVFGATKG 720
 KRFPEDLAQQ NLIISNTEAP GDDRVCSSANG FMTQTNNSS QGFCGTMGSG MKNGGQETIE 780
 MMKGNGQTL SCRGAGHHT LDSCRGGHTE VDNCRYTSE WHSFTQPLRG EKLHRCNQNE 840
 35 DRMPSQDYVL TYNYEGRGSP AGSVGCCSEK QBEDGLDPLN NLEPKFITLA EACTKR

40 Seq ID NO: 530 DNA sequence
 Nucleic Acid Accession #: NM_016583.2
 Coding sequence: 72..842

1 11 21 31 41 51
 45 GGAGTGGGGG AGAGAGAGGA GACCAGGACA GCTGCTGAGA CCTCTAAGAA GTCCAGATAC 60
 TAAGAGCAAA GATGTTTCAA ACTGGGGGCC TCATGTCTCT CTACGGGCTG TTAGCCCAGA 120
 CCATGGCCCA GTTTGGAGGC CTGCCGTGTC CCCTGGACCA GACCCGTGCC TTGAATGTGA 180
 ATCCAGCCCT GCCCTTAGT CCCACAGGTC TTGCAGGAAG CTTGACAAAT GCCCTCAGCA 240
 ATGGCCTGCT GTCTGGGGGC CTGTTGGGCA TTCTGGAAA CCTTCCGCTC CTGGACATCC 300
 TGAAGCCTGG AGGAGGTACT TCTGGTGGCC TCCTTGGGGG ACTGCTTGA AAAGTGACGT 360
 50 CAGTGATTCC TGGCCTGAAC AACATCATG ACATAAAGGT CACTGACCCC CAGCTGCTGG 420
 AACTTGGCCT TGTGCAGAGC CCTGATGGCC ACCGTCTCTA TGTCAACATC CCTCTCGCA 480
 TAAAGCTCCA AGTGAATACG CCCCTGGTCG GTGCAAGTCT GTTGAGGCTG CTRGTGAAGC 540
 TGGACATCAC TGCAAAATC TTAGCTGTGA GAGATAAGCA GGAGAGGATC CACCTGGTCC 600
 TTGTTGACTG CACCCATTCC CTTGGAAGCC TGCAAAATTC TCTGCTGAT GGACTTGGCC 660
 55 CCTCTCCCAT TCAAGTCTT CTGACAGGCC TCACAGGGAT CTGATAAA GTCTCTCTG 720
 AGTTGGTTCA GGGCAACGTG TGCCCTCTGG TCAATGAGT TCTCAGAGC TTGACATCA 780
 CCTGTGTGCA TGACATTGTT AACATGCTGA TCCACGGACT ACAGTTTGTG ATCAAGSTCT 840
 AAGCCTTCCA GGAAGGGGCT GGCCCTCTGT GAGCTGCTC CCACTGTCTA CAGATGGCTG 900
 GCCCATGTGC TGAAGATGCA CACAGTTGCC TTCTCTCCGA GGAACCTGCC CCTCTCTCT 960
 60 TCCCACAGG CGTGTGTAAC ATCCCATGTG CCTCACCTAA TAAATGGCT CTCTCTCTG 1020
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

65 Seq ID NO: 531 Protein sequence
 Protein Accession #: NP_057667.1

1 11 21 31 41 51
 70 MFQTGGLIVF YGLLAQTMAQ FGGLPVPLDQ TLPLNVNPAI PLSPTGLAGS LTNALSNGLL 60
 SGGLLGILEN LPLLDILRPG GGTSGLLGG LLGKVTSVIP GLNNIIDIKV TDPQLLELGL 120
 VQSPDGHRLY VTPIPLGIKQ VNTPLVGASL LRLAVKLDIT AEILAVRDQK ERIHLVLGDC 180
 75 TLSPGSLQIS LLDGLGPLFI QGLLDSLGTI LNKVLPVLVQ GNVCPLVNEV LRGLDITLVH 240
 DIVNMLIHGL QFVIKV

80 Seq ID NO: 532 DNA sequence
 Nucleic Acid Accession #: NM_004363.1
 Coding sequence: 115..2223

1 11 21 31 41 51
 80 CTCAGGGCAG AGGGAGGAAG GACAGCAGAC CAGACAGTCA CAGCAGCCTT GACAAAACGT 60
 TCTTGGAACT CAAGCTCTTC TCCACAGAGG AGGACAGAGC AGACAGCAGA GACCATGGAG 120
 TCTCCCTCGG CCCCTCCCA CAGATGGTGC ATCCCTGGC AGAGGCTCTT GCTCAGAGCC 180
 TCACTTCTAA CCTTCTGGAA CCGGCCACCC ACTGCCAAGC TCACTATTGA ATCCAGCGCG 240
 85 TTCAATGTGG CAGAGGGGAA GGAGGTGCTT CTACTGTGCC ACAATCTGCC CCAGCATCTT 300
 TTTGGCTACA GCTGGTACAA AGGTGAAAGA GTGGATGGCA ACCGTCAAAT TATAGGATAT 360
 GTAATAGGAA CTCACAAGC TACCCAGGG CCGCATACA GTGGTCGAGA GATAATATAC 420
 CCCAATGCAT CCTGCTGAT CCAGAACATC ATCCAGAATG ACACAGGATT CTACACCTTA 480

5
10
15
20
25
30
35
40
45
50
55
60

Seq ID NO: 533 Protein sequence
Protein Accession #: NP_004354.1

1 11 21 31 41 51
MESPSAPPHR WCIPWQRLLL TASLLTFWNP PTTAKLTIES TPFNVAEKKE VLLLVHNLPO 60
HLFGYSWYKG ERVDGNRQII GYVIGTQQAT PGPAYSGREI IYPNASLLIQ NIIQNDTGFY 120
TLHVIKSLDLV NEEATGQFRV YPELPKPSIS SNNSKPVEDK DAVAFTEPE TQDATYLLWWV 180
NNQSLVSPR LQLSNGRNL TLFNVTRNDT ASYKCTQNP VSARRSDSVI LNVLYGPDAP 240
TISPLNTSYR SGENLNLSC AASNPPAQYS WFNVTGFOQS TQELFIPNIT VMNSSGYTCQ 300
AHNSDTGLNR TTVITITVYA EPPKPFITSN NSNPVEDEDA VALTCEPEIQ NTTYLWVNN 360
QSLVPSPRLQ LSNDRNLTL LSVTRNDVGP YECGIQNELS VDHSDPVILN VLYGPDPTI 420
SPSYTYRPG VNLSLSCHAA SNPPAQYSWL IDGNIQHTQ ELFIENITEK NSGLYTCQAN 480
NSAGHSRRT VKTITVSABL PKPSISSNNS KPEVEDKAVA FTCEPEAQNT TYLWVWNGQS 540
LPVSPRLQLS NGNRTLTLFN VTRNDARAYV CGIQNSVSAN RSDPVTLDVL YGPDPTIISP 600
PDSSYLSGAN LNLSCHSASN PSPQYSWRIN GIPQHTQVL FIAKITPNVN GYACFVSNL 660
ATGRNNSIVK SITVSAGTS PGLSAGATVG IMIGVLGVVA LI

Seq ID NO: 534 DNA sequence
Nucleic Acid Accession #: NM_006952.1
Coding sequence: 11..793

1 11 21 31 41 51
AATCCCGACA ATGGCGAAAG ACAACTCAAC TGTTCGTGTC TTCCAGGGCC TGCTGATTTT 60
TGGAAATGTG ATTATTGGTT GTTGCGGCAT TGCCCTGACT GCGGAGTGCA TCTTCTTTGT 120
ATCTGACCAA CACAGCCTCT ACCCACTGCT TGAAGCCACC GACAACGATG ACATCTATGG 180
GGCTGCGCTG ATCGGCATAT TTGTGGGCAT CTGCCTCTTC TGCCGTGCTG TCTTAGGCAT 240
TGTAGGCATC ATGAAGTCCA GCAGGAAAAT TCTTCTGGCG TATTTCATTC TGATGTTTAT 300
AGTATATGCC TTGAAGTGG CATCTGTAT CACAGCAGCA ACACAACGAG ACTTTTTCAC 360
ACCCAACTCT TTCTGAAGC AGATGCTAGA GAGGTACCAA AACACAGCC CTCCAACAA 420
TGATGACCA TGGAAAAACA ATGGAGTCAC CAAAACCTGG GACAGGCTCA TGCTCCAGGA 480
CAATTGCTGT GCGGTAAATG GTCCATCAGA CTGGCAAAAA TACACATCTG CTTTCCGGAC 540
TGAGAATAAT GATGCTGACT ATCCCTGGCC TGTCAATGTC TGTGTTATGA ACAATCTTAA 600
AGAACCTCTC AACCTGGAGG CTGTGAAACT AGGCGTGCTT GGTTTTATC ACAATCAGGG 660
CTGCTATGAA CTGATCTCTG GTCCAATGAA CCGACACGCC TGGGGGGTTG CTTGTTTGG 720
ATTTGCCATT CTCTGCTGGA CTTTGTGGGT TCTCTGGGT ACCATGTTCT ACTGGAGCAG 780
AATTGAATAT TAAGAA

Seq ID NO: 535 Protein sequence
Protein Accession #: NP_008883.1

1 11 21 31 41 51
AATCCCGACA ATGGCGAAAG ACAACTCAAC TGTTCGTGTC TTCCAGGGCC TGCTGATTTT 60
TGGAAATGTG ATTATTGGTT GTTGCGGCAT TGCCCTGACT GCGGAGTGCA TCTTCTTTGT 120
ATCTGACCAA CACAGCCTCT ACCCACTGCT TGAAGCCACC GACAACGATG ACATCTATGG 180
GGCTGCGCTG ATCGGCATAT TTGTGGGCAT CTGCCTCTTC TGCCGTGCTG TCTTAGGCAT 240
TGTAGGCATC ATGAAGTCCA GCAGGAAAAT TCTTCTGGCG TATTTCATTC TGATGTTTAT 300
AGTATATGCC TTGAAGTGG CATCTGTAT CACAGCAGCA ACACAACGAG ACTTTTTCAC 360
ACCCAACTCT TTCTGAAGC AGATGCTAGA GAGGTACCAA AACACAGCC CTCCAACAA 420
TGATGACCA TGGAAAAACA ATGGAGTCAC CAAAACCTGG GACAGGCTCA TGCTCCAGGA 480
CAATTGCTGT GCGGTAAATG GTCCATCAGA CTGGCAAAAA TACACATCTG CTTTCCGGAC 540
TGAGAATAAT GATGCTGACT ATCCCTGGCC TGTCAATGTC TGTGTTATGA ACAATCTTAA 600
AGAACCTCTC AACCTGGAGG CTGTGAAACT AGGCGTGCTT GGTTTTATC ACAATCAGGG 660
CTGCTATGAA CTGATCTCTG GTCCAATGAA CCGACACGCC TGGGGGGTTG CTTGTTTGG 720
ATTTGCCATT CTCTGCTGGA CTTTGTGGGT TCTCTGGGT ACCATGTTCT ACTGGAGCAG 780
AATTGAATAT TAAGAA

WO 02/086443

PCT/US02/12476

MAKDNSTVRC FQGLIFGNV IIGCCGIALT AECIFFVSDQ HSLYPLLEAT DNDDIYGAAW 60
 IGIFVGLICLF CLSVLIGIVI MKSSRKILLA YFILMFIVYA FEVASCITAA TQRDFFTPNL 120
 FLKQMLERVQ NNSPPNDDQ WKNNVGTWV DRLMLQDNCC GVNGPSDWQK YTSAFRTENN 180
 DADYFWPRQC CVMNLEKPL NLEACKLGPV GFYHNGQCYE LISGPMNRHA WGVANWFGFAI 240
 LCWTFWVLLG TMFYWSRIEY

Seq ID NO: 536 DNA sequence
 Nucleic Acid Accession #: NM_002638.1
 Coding sequence: 120..473

1 11 21 31 41 51
 CAATACAGCT AAGGAATTAT CCCTTGTAAT TACCACAGAC CGGCCCTGGA GCCAGGCCAA 60
 GCTGGAGCTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120
 TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCTCAT CGCTGGGACG CTGGTCTTAG 180
 AGGCAGCTGT CACGGGAGTT CCTGTTAAAG GTCAGACAC TGTCAAAGGC CGTGTCTCAT 240
 TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAGTTAA AGGTCAAGAT AAGTCAAAG 300
 CGCAAGAGCC AGTCAAGGT CCAGTCTCCA CTAAGCTGG CTCCTGCCCC ATTATCTTGA 360
 TCCGGTGGCG CATGTTGAAT CCCCCTAACC GCTGCTTGA AGATACTGAC TGCCAGGAA 420
 TCAAGAAGTG CTGTGAAGGC TCTTGGCGGA TGGCCTGTT CGTTCCCGAG TGAAGGAGC 480
 CGGTCTTGG TGCACCTGTG CCGTCCCGAG AGCTACAGGC CCCATCTGGT CCTAAGTCCC 540
 TGCTGCCCTT CCCCCTTCCA CACTGTCCAT TCTTCTCCC ATTCAAGATG CCCAGGCTG 600
 GAGCTGCTC TCTCATCCAC TTTCATAATA A

Seq ID NO: 537 Protein sequence
 Protein Accession #: NP_002629.1

1 11 21 31 41 51
 MRASSPLIVV VFLIAGTLVL EAAVTGVPVK GQDTVKGRVP FNGQDPVKQ VSVKQGDVKV 60
 AQEPVKGPIV TKPGSCPIIL IRCAMLNPPN RCLKDTDCPG IKKCEGSCG MACFVPG

Seq ID NO: 538 DNA sequence
 Nucleic Acid Accession #: NM_001793.2
 Coding sequence: 71..2560

1 11 21 31 41 51
 AAAGGGCAA GAGCTGAGCG GAACACCGGC CGGCCGTGCG GGCAGCTGCT TCAACCCCTCT 60
 CTCTGCGACC ATGGGGCTCC CTGCTGGACC TCTCGCTCT CTCTCTCTT TCCAGGTTTG 120
 CTGGCTCAG TGCGCGCGCT CCGAGCCGTG CCGGGCGGTC TTCAGGAGG CTGAAGTGAC 180
 CTGGAGGCG GGAGGCGCGG AGCAGGAGCC CGGCCAGCG CTGGGAAAG TATTCATGGG 240
 CTGCCCTGGG CAAGAGCCAG CTCTGTTTAG CACTGATAAT GATGACTTCA CTGTGCGGAA 300
 TGGCGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAAG AATCCATTGA AGATCTTCCC 360
 ATCCAAAGCT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCAA TATCTGTCCC 420
 TGAATATGGC AAGGTCCTCT TCCCCAGAG ACTGAATCAG CTCAAGTCTA ATAAAGATAG 480
 AGACACCAAG ATTTTCTACA GCATCACGGG GCCGGGGGCA GACAGCCCCC CTGAGGCTGT 540
 CTTCGCTGTA GAGAGGAGA CAGGCTGGTT GTTGTGTAAT AAGCCACTGG ACCGGGAGGA 600
 GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTCAGAG AATGGTGCCT CAGTGGAGGA 660
 CCCCATGAAC ATCTCCATCA TCGTGACCGA CCAGAATGAC CACAAGCCCA AGTTTACCCA 720
 GGACACCTTC CGAGGGAGTG TCTTAGAGGG AGTCTTACCA GGTACTTCTG TGATGCAGGT 780
 GACAGCCACG GATGAGGATG ATGCCATCTA CACCTACAAT GGGGTGGTTG CTTACTCCAT 840
 CCATAGCCAA GAACCAAGG ACCCACAGA CCTCATGTTT ACCATTACCC GGAGCAGG 900
 CACCATCAG GTCACTCCA GTGGCCTGGA CCGGGAAAAA GTCCCTGAGT ACACACTGAC 960
 CATCCAGGCC ACAGACATGG ATGGGGACGG CTCCACACCC ACGGCAGTGG CAGTATGTTG 1020
 GATCTCTGAT GCCAATGACA ATGCTCCCAT GTTGAACCCC CAGAAGTACG AGGCCCATGT 1080
 GCCTGAGAAAT GCATGGGGCC ATGAGGTGCA GAGGCTGACG GTCACTGATC TGGACGCCCC 1140
 CAACCTACCA CGGTGGCGTG CCACCTACCT TATCATGGGC GGTGACGACC GGGACCATTT 1200
 TACCATCACC ACCCAACCTG AGAGCAACCA GGGCATCCTG ACACCCAGGA AGGGTTTGA 1260
 TTTTGAAGCC AAAAACACG ACACCTGTA CGTTGAAGTG ACCAAGAGG CCCCTTTTGT 1320
 GCTGAGGCTC CCAACCTCCA CAGCCACCAT AGTGGTCCAG GTGGAGGATG TGAATGAGGC 1380
 ACCTGTGTTT GTCCCAACCT CCAAAGTCGT TGAGGTCCAG GAGGGCATCC CCACTGGGGA 1440
 GCCTGTGTGT GTCTACACTG CAGAAGACCC TGACAAGGAG AATCAAAAGA TCGACTACCG 1500
 CATCTGTAGA GACCCAGCAG GGTGGCTAGC CATGGACCCA GACAGTGGGC AGGTACACGC 1560
 TGTGGGCACC CTGACCGGTG AGGATGAGCA GTTGTGAGG AACAACTCT ATGAAGTCAT 1620
 GGTCTTGGCC ATGGACAATT GAAGCCCTCC CACCACTGGC ACGGGAACCC TTCTGCTAAC 1680
 ACTGATTGAT GTCAATGACC ATGGCCCACT CCCTGAGCCC CGTCAGATCA CCATCTGCAA 1740
 CCAAAGCCCT GTGCGCCAGG TGCTGAACAT CACGGACAAG GACCTGTCTC CCCACACCTC 1800
 CCCTTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGGACGGCAG AGGTCAACGA 1860
 GGAAGGTGAC ACAGTGGTCT TGTCCCTGAA GAAGTTCCTG AAGCAGGATA CATATGACGT 1920
 GCACCTTTCT CTGCTGACCC ATGSCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980
 GTGCGACTGC CATGGCCATG TCGAAACCTG CCCTGGACCC TGAAGGGAG GTTTCATCCT 2040
 CCCTGTGCTG GGGGCTGTCC TGGCTCTGCT GTTCTCTCTG CTGGTGTGCT TTTTGTGTTG 2100
 GAGAAAGAG CGGAAGATCA AGGAGCCCCC CTAATCTCCA GAAGATGACA CCGGTGACAA 2160
 CGTCTTCTAC TATGCGCAAG AGGGGGGTGG CGAAGAGGAC CAGGACTATG ACATCACCCA 2220
 GCTCCACCGA GGTCTGGAGG CCAGGCGGGA GGTGGTCTC CGCAATGAGC TGGCACCAAC 2280
 CATCATCCCG ACACCCATGT ACCGTCTCTG GCCAGCCAAC CCAGATGAAA TOGGCAACTT 2340
 TATAATTGAG AACCTGAAGG CGGCTAACAC AGACCCCAAC GCCCGCCCTC ACCACACCTT 2400
 CTGTGTGTTT GACTATGAGG GCAAGCGGCTC CGACGCGGCT TCCCTGAGCT CCTCACCTC 2460
 CTCGCCCTCC GACCAAGACC AAGATTACGA TTATCTGAAC GAGTGGGGCA GCCGCTTCAA 2520
 GAAGCTGGCA GACATGTACG GTGGCGGGGA GGACGACTAG GCGGCTGCTC TGCAGGCTG 2580
 GGGACCAAC GTGAGCCAC AGAGCATCTC CAAGGGGTCT CAGTTCCTCC TTCAGCTGAG 2640
 GACTTCGAG GTGTGACGGA AGTGGCCGTA GCAACTTGGC GGAGACAGGC TATGAGTCTG 2700
 ACGTTAGAGT GGTGTCTTCC TTAGCCTTTC AGGATGGAGG AATGTGGGCA GTTGTGACTT 2760
 AGCACTGAAA ACCTCTCCAC CTGGGCCAGG GTTGCCTCAG AGGCCAAGTT TCCAGAAGCC 2820
 TCTTACCTGC CGTAAATGTC TCAACCTGTG TTCTAGGCC TGGGCTGCT GTGACTGACC 2880
 TACAGTGGAC TTTCTCTCTG GAATGGAACC GTCTTAGGCC TCTGTGTGCA ACTTAATTTT 2940

TTTTTTTAAAT GCTATCTTCA AAACGTTAGA GAAAGTTCTT CAAAAGTGCA GCCCAGAGCT 3000
 GCTGGGCCCA CTGGCCGTCC TGCATTCTG GTTTCAGAC CCCAATGCCT CCCATTCCGA 3060
 TGGATCTCTG CGTTTTTATA CTGAGTGTGC CTAGTTGCC CCTATTITTT TATTTTCCCT 3120
 GTTCCGTGCG TATAGATGAA GGGTGAGGAC AATCGTGTAT ATGTACTAGA ACTTTTTTAT 3180
 TAAAGAACT TTTCCAGAA AAAAA

Seq ID NO: 539 Protein sequence
 Protein Accession #: NP_001784.2

1 11 21 31 41 51
 MGLPRGPLAS LLLQVCWLQ CAASEPCRAV FREAEVTLEA GGAQEPEGQA LGKVPMGCPG 60
 QEPALFSTDN DDFTVRNGET VQERRSLKER NPLKIPPSKR ILRRHKRDWV VAPISVPENG 120
 KGPPFQRLNQ LKSNKDRDTK IFYSITGPGA DSPPEGVFAV EKETGWLLLN KPLDREEIAK 180
 YELFGHAVSE NGASVEDPMN ISIIVTDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT 240
 DEDDAIYVTN GVVAISIHSQ EFKDPHDLMP TIHRSTGTIS VISSGLDREK VPEYTLTIQA 300
 TDMGDGOSTL TAVAVVILDL ANDNAPMFDP QKYEAHVPEN AVGHEVQRLT VTDLDAPNSP 360
 ANRATYILMG GDDGDHFTIT THPESNQIIL TTRKGLDFEA KNQHTLYVEV TNEAFPVLKL 420
 PTSTATIVVH VEDVNEAPVE VPPSKVVEVQ EGIPTGEPVC VYTAEDPDKE NQKISYRILR 480
 DPAGWLAMD PDSGQVAVGT LDREDEQFVR NNIYEVMLA MDNGSPPTTG TGTLLLTLLD 540
 VNDHGPPVEP RQITICNQSP VRQVLNITDK DLSPTSPFQ AQLTDDSDIY WTAEVNNEGD 600
 TVVLSLKKFL KQDTYDVHLS LSDHGNKEQL TVIRATVDCD HGHVETCPGP WKGGFILPVL 660
 GAVLALLFL LVLVLLVRKK RKIKEPLLLP EDDTRDNVFI YGEBGGGEED QDYDITQLHR 720
 GLEARPEVVL RNDVAPTIIIP TFMYPRPFPAN PDEIGNFIE NLKAANTDPT APPYDTLLVF 780
 DYESGSGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD

Seq ID NO: 540 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..672

1 11 21 31 41 51
 ATGAGGCTCC AAAGACCCCG ACAGGCCCGG CGGGGTGGGA GCGCGCGGCC CGGGGGCGGG 60
 CGGGGCTCCC CCTACGGGCC AGACCCGGGG AGAGGCGCGC GGAGGCTGCG AAGGTTCCAG 120
 AAGGGCGGGG AGGGGGCGCC GCGCGCTGAC CCTCCCTGGG CACCGCTGGG GACGATGGCG 180
 CTGCTCGCCT TGCTGCTGGT CGTGGCCCTA CCGCGGGTGT GGACAGACGC CAACCTGACT 240
 GCGAGACAAC GAGATCCAGA GGAATCCAGC CGAACGGAGC AGGGTGACAA TAGAGTGTGG 300
 TGTGATGTTT GTGAGAGAGA AAACACTTTC GAGTGCCAGA ACCCAAGGAG GTGCAATGG 360
 ACAGAGCCAT ACTGCGTTAT AGCGGCGGTG AAAATATTTT CACGTTTTTT CATGGTTGCG 420
 AAGCAGTGTG CCGCTGTTTG TGCAGCGATG GAGAGACCCA AGCCAGAGGA GAAGCGGTTT 480
 CTCTGTGAAG AGCCCATGCC CTTCTTTTAC CTCAGTGTGT GTAAATTCG CTACTGCAAT 540
 TTAGAGGGGC CACCTATCAA CTCATCAGTG TTCAAAGAAT ATGCTGGGAG CATGGGTGAG 600
 AGCTGTGGTG GGTGTGGTGT GGCCATCCTC CTGCTGCTGG CCTCAATTGC AGCGCGCCTC 660
 AGCCTGTCTT GA

Seq ID NO: 541 Protein sequence
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 MRLQRPRQAP AGRRAPRGG RGSPYRPDPG RGARRLRFP KGGEGAPRAD PPWAPLGTMA 60
 LLALLLVVAL PRVWTDANLT ARQRDPEDSQ RTDEGNRVW CHVCHERNTF EQCNPRRCKW 120
 TEPYCVIAAV KIFPRPFMVA KQCSAGCAAM ERPKPEEKRF LLEPMPPFFY LKCCKIRYCN 180
 LEGPFINSSV FKEYAGSMGE SCGGLWLAIL LLLASIAAGL SLS

Seq ID NO: 542 DNA sequence
 Nucleic Acid Accession #: XM_035292.2
 Coding sequence: 53..1576

1 11 21 31 41 51
 GCTCGCTGGG CCGCGGCTCC CGGGTGTCCC AGGCCCGGCC GGTGCGCAGA GCATGGCGGG 60
 TCGGGGCGCG AAGCGGCGCG CGCTAGCGGC GCCGGCGGCC GAGGAGAAGG AAGAGGCGCG 120
 GGAGAAGATG CTGGCCGCCA AGAGCGCGGA CGGCTCGGCG CCGGAGGCG AGGGCGAGGG 180
 CGTGACCCCTG CAGCGGAACA TCACGCTGCT CAACGGCGTG GCCATCATCG TGGGGACCAT 240
 TATCGGCTCG GGCATCTTCG TGACGCCAC CCGCGTGTCT AAGGAGGAG GCTCGCGGG 300
 GCTGGCGCTG GTGGTGTGGG CCGCGTGTGG CGTCTTCTCC ATCGTGGGCG CGCTCTGCTA 360
 CGCGGAGCTC GGCACCAACA TCTCCAAATC GGGCGGCGAC TACGCTACA TGCTGGAGGT 420
 CTACGGCTCG CTGCGCGCCT TCCTCAAGCT CTGGATCGAG CTGCTCATCA TCGGCGCTTC 480
 ATCGCAGTAC ATCGTGGGCC TGCTCTTCGC CACCTACCTG CTCAGGCGCG TCTTCCCCAC 540
 CTGCGCGGTG CCGGAGGAGG CAGCCAAAGT CGTGGCGTGC CTCTGCTGTC TGCTGCTCAC 600
 GGCCTGGAAC TGCTACAGCG TGAAGGCCGC CACCGGGTTC CAGGATGCCT TTGCGCGCGC 660
 CAAGCTCCTG GCCCTGGCCC TGATCATCCT GCTGGGCTTC GTCCAGATCG GAAAGGGTGA 720
 TGTGTCCAAAT CTAGATCCCA ACTTCTCAT TGAAGGCACC AAACCTGGATG TGGGGAACAT 780
 TGTGTGGCA TTATACAGCG GCCTCTTTCG CTATGGAGGA TGAATTACT TGAATTGCT 840
 CACAGAGGAA ATGATCAACC CCTACAGAAA CCGTCCCGTG GCCATCATCA TCTCCCTGCC 900
 CATCGTGACG CTGGTGTAGC TGCTGACCAA CCGGCTTACC TTCAACACCC TGTCCACCGA 960
 GCAGATGCTG TCGTCCGAGG CCGTGGCGGT GGAATTCGGG AACTATCACC TGGGCGTCAT 1020
 GTCTTGGATC ATCCCGTCTT TCGTGGGCGT GTCTTCTGTT CGCTCCGTC AATGGTCCCT 1080
 GTTCAATCC TCCAGGCTCT TCTTCTGTGG GTCCCGGGAA GGCCACCTGC CCTCCATCCT 1140
 CTCCATGATC CACCCAGAGC TCCTCACCCC CGTGGCGTCC CTGCTGTCA CGTGTGTGAT 1200
 GACGCTGCTC TAGCCTTCTT CCAAGGACAT CTTCTCCGTC ATCAACTTCT TCAGCTTCTT 1260
 CAACTGGCTC TGGTGGGCC TGGCCATCAT CGGCATGATC TGGCTCGGCC ACAGAAAGCC 1320
 TGAGCTTGA AGGCCCATCA AGGTGAACCT GGCCTTGCTT GTGTTCTTCA TCTTGGCCTG 1380
 CCTCTTCTG ATCGCGTCTC CTTTCTGAA GACACCGGTG GAGTGTGGCA TCGGCTTCC 1440
 CATCATCTC AGCGGCTGC CGTCTACTT CTTGCGGGTC TGGTGAAGAA ACAAGCCCAA 1500
 GTGGCTCTC CAGGCGATCT TCTCCACGAC CGTCTGTGT CAGAAGCTCA TGCAGTGTGT 1560

CCCCCAGGAG ACATAGCCAG GAGGCCGAGT GGCTGCCGGA GGAGCATGC

Seq ID NO: 543 Protein sequence
Protein Accession #: XP_035292.2

5
10
15
20

1	11	21	31	41	51	
MAGAGPKRRA	LAAPAAEEKE	EAREKMLAAK	SADGSAPAGE	GEGVTIQRNI	TLINGVAIIV	60
GTIIGSGIFV	TPTGVLKEAG	SPGLALVWVA	ACGVFSIVGA	LCYAEIGTTI	SKSGGDYAYM	120
LEVYGSIPAF	LKLWIELLII	RPSSQYIVAL	VFATYLLKPL	FPTCPVPEEA	AKLVACLCLV	180
LLTAVNCYSV	KAATRVQDAF	AAAKLLALAL	IILLGFVQIG	KGDVSNLDPN	FSFEGTKLDV	240
GNIVLALYSY	LFAYGWNVYL	NEVTEEMINP	YRNLPLAIII	SLPIVTLVYV	LINLAYFTTL	300
STEQMLSSSA	VAVDFGNVHL	GVMSHIIPVF	VGLSCFGSVN	GSFLTSSRLF	FVGSREGHLP	360
SLSMIHPQL	LTPVPSLVPT	CVMTLLYAFS	KDIFSINFPF	SFFNWLCLAL	AIIGMIWLRH	420
RKPELERPIK	VNLALPVFFI	LACLFLIAVS	PWKTPVECGI	GFTIILSGLP	VYFFGVWVKN	480
KPKNLLQGIF	STTVLCQKLM	QVVPQET				

Seq ID NO: 544 DNA sequence
Nucleic Acid Accession #: NM_005268.1
Coding sequence: 168..989

25
30
35
40
45

1	11	21	31	41	51	
TAAAAAGCAA	AAGAATTCGC	GGCCGCGTCG	ACACGGGCTT	CCCCGAAAAC	CTTCCCCGCT	60
TCTGGATATG	AAATTCAGCG	TGCTTGCTGA	GTCCTATTGC	CGGCTGTGGG	GAGCCAGGAG	120
AGCCCTGAGG	AGTAGTCACT	CAGTAGCAGC	TGACGCGTGG	GTCCACCATG	AACTGGAGTA	180
TCTTTGAGGG	ACTCTGAGT	GGGCTCAACA	AGTACTCCAC	AGCCTTTGGG	CGCATCTGGC	240
TGCTCTGGT	CTTCATCTTC	CGGCTGTGGG	TGTACTCTGT	GACGCGCCAG	CGTGTGTGGA	300
GTGATGACCA	CAAGGACTTC	GACTGCAATA	CTGCCAGGCC	CGGCTGTGCC	AACGCTGTCT	360
TGATGAGTT	CTTCCCTGTG	TCCCATGTGC	GCCTCTGGGC	CCTGCAGCTT	ATCCTGGTGA	420
CATGCCCTTC	ACTGCTCGTG	GTCATGCACG	TGGCCTACCG	GGAGGTTTCA	GAGAAGAGGC	480
ACCGAGAAGC	CACTGGGGAG	AACAGTGGGC	GCCTCTACCT	GAACCCCGGC	AAGAAGCGGG	540
GTGGGCTCTG	GTGGACATAT	GTCTGCAGCC	TAGTGTTCAC	GGCGAGCGTG	GACATGCGCT	600
TTCTCTATGT	GTTCACACTCA	TTCTACCCCA	AATATATCCT	CCCTCTCTGT	GTCAAGTGCC	660
ACGCAGATCC	ATGTCCCAAT	ATAGTGGACT	GCTTCATCTC	CAAGCCCTCA	GAGAAGAACA	720
TTTTCACTCT	CTTCATGCTG	GCCACAGCTG	CCATCTGCAT	CCTGCTCAAC	CTCGTGAGC	780
TCATCTACCT	GGTAGCAAG	AGATGCCACG	AGTGCCTGGC	AGCAAGGAAA	GCTCAAGCCA	840
TGTGCACAGG	TCATCACCCC	CACGCTACCA	CTCTCTCTGT	CAACAAGAC	GACCTCTTTT	900
CGGGTGACCT	CATCTTCTGT	GGCTCAGACA	GTCTCTCTCT	TCTCTTACCA	GACCGCCCCC	960
GAGACCATGT	GAGAAAACCC	ATCTTGTGAG	GGGCTGCGCT	GACTGTGCTG	GCAGGTGTGG	1020
CCTGGATGGG	GAGGCTCTAG	CATCTCTCAT	AGGTGCAACC	TGAGAGTGGG	GGAGCTAAGC	1080
CATGAGGTAG	GGGAGGCCAA	GAGAGAGGAT	TCAGACGCTC	TGGGAGCCAG	TTCTAGTCC	1140
TCAACTCCAG	CCACTGCGCC	CAGCTCGACG	GCATGGGGCC	AGTTCCTCC	CTGCTCTGCA	1200
GCTGGTTTTC	CTTTTCTAGA	ATGGAAATAG	TGAGGGCCAA	TGC		

Seq ID NO: 545 Protein sequence
Protein Accession #: NP_005259.1

50
55

1	11	21	31	41	51	
MNWSIFEGLL	SGVNKYSTAF	GRWLSLVFI	FRVLVYLVT	ERVWSDHKD	FDCNTRQPGC	60
SNVCFDEFFP	VSHVRLMALQ	LILVTCPSLL	VVMHVAYREV	QEKRHREAHG	ENSGRLYLNP	120
GKKRGGLWMT	YVCSLVFKAS	VDIAFLYVFH	SFYPKYILPP	VVKCHADPCP	NIUDCFISKP	180
SEKNIFTLFM	VATAAICILL	NLVELIYLV	KRCHECLAAR	KAQAMCTGHH	PHGTTSSCKQ	240
DDLSSGDLIF	LGSDSRPPLL	PDRPRDHVK	TIL			

Seq ID NO: 546 DNA sequence
Nucleic Acid Accession #: NM_002391.1
Coding sequence: 26..457

60
65
70
75

1	11	21	31	41	51	
CGGGCGAAGC	AGCGCGGGCA	GCGAGATGCA	GCACCGAGGC	TTCTCTCTCC	TCACCTCCTC	60
CGCCCTGCTG	GCGCTCACCT	CCGCGGTCCG	CAAAAAGAAA	GATAAGGTGA	AGAAGGGCGG	120
CCCGGGGAGC	GAGTGCCTGT	AGTGGGCTGT	GGGGCCCTGC	ACCCCCAGCA	GCAAGGATTG	180
CGCGCTGGGT	TTCGCGAGG	GCACCTGCGG	GGCCAGACCC	CAGCGCATCC	GGTGCAGGGT	240
CGCCTGCAAC	TGAAGAAGG	AGTTTGGAGC	CGACTGCAAG	TACAAGTTTG	AGAAGTGGGG	300
TGCGTGTGAT	GGGGGACAG	GCACCAAGT	CCGCCAAGGC	ACCTGGAAGA	AGGCGCGCTA	360
CAATGCTCAG	TGCCAGGAGA	CCATCCGCGT	CACCAAGCCC	TGCACCCCCA	AGACCAAGGC	420
AAAGGCCCAA	GCCAGGAAG	GGAAGGGAAG	GGACTAGACG	CCAAGCCTGG	ATGCCAAGGA	480
GCCCCCTGGT	TCACATGGGG	CCTGGCCACG	CCCTCCCTCT	CCCAGGCCCG	AGATGTGACC	540
CACCATGCCC	TTCTGTCTGC	TCGTTAGCTT	TAATCAATCA	TGCCCTGACT	TGTCCCTCTC	600
ACTCCCCAGC	CCCCACCTTA	AGTGCCCAAA	GTGGGGAGGG	ACAAGGGATT	CTGGGAAGCT	660
TGAGCTCCCC	CCAAAGCAAT	GTGAGTCCCA	GAGCCCGCTT	TTGTTCTTCC	CCACAATTCC	720
ATTACTAAGA	AACACATCAA	ATAAACTGAC	TTTTTCCCCC	CAATAAAGC	TCTTCTTTTT	780
TAATAT						

Seq ID NO: 547 Protein sequence
Protein Accession #: NP_002382.1

80
85

1	11	21	31	41	51	
MQHRGFLLLT	LLALLALTS	VAKKKDKVK	GGPGSECAEW	AWGPCTPSSK	DCGVGFREGT	60
CQAQTQIRIC	RVPCNWKKEF	GADCKYKFEN	WGACDGGTGT	KVRQGLTKKA	RYNAQCQETI	120
RVTKPCTPKT	KAKAKAKGK	GKD				

Seq ID NO: 548 DNA sequence

Nucleic Acid Accession #: NM_006783.1
Coding sequence: 1..786

```
5 1 11 21 31 41 51
ATGGATTGGG GGACGCTGCA CACTTTCATC GGGGGTGTCA ACAAACACTC CACCAGCATC 60
GGGAAGGTGT GGATCAGAGT CATCTTTATT TTCCGAGTCA TGATCCTAGT GGTGGCTGCC 120
CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTGCTCTGCA ACACACTGCA ACCGGGATGC 180
AAAAATGTGT GCTATGACCA CTTTTTCCCG GTGTCCACCA TCCGGCTGTG GGCCCTCCAG 240
10 CTGATCTTCG TCTCCACCCC AGCGCTGCTG GTGGCCATGC ATGTGGCCTA CTACAGGCAC 300
GAAACCACTC GCAAGTTTCA GCGAGGAGAG AAGAGGAATG ATTTCAAAGA CATAGAGGAC 360
ATTAAAAAGC ACAAGTTTCG GATAGAGGGG TCGCTGTGGT GGACGTACAC CAGCAGCATC 420
TTTTTCCGAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTTACTTCCT TTACAATGGG 480
TACCACCTGC CTGGGTGTTT GAAATGTGGG ATTGACCCCT GCCCAACCT TGTGTACTGC 540
15 TTTATTCTTA GGCCAAACAGA GAAGACCGTG TTTACCATTT TTATGATTTC TGCGTCTGTG 600
ATTGTCATGC TGCTTAACGT GGCAGAGTTG TGCTACCTGC TGCTGAAAGT GTGTTTTAGG 660
AGATCAAAGA GAGCACAGAC GCAAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG 720
CAGAATGAAA TGAATGAGCT GATTTCAGAT AGTGGTCAAA ATGCAATCAC AGGTTTCCCA 780
AGCTAA
```

Seq ID NO: 549 Protein sequence
Protein Accession #: NP_006774.1

```
25 1 11 21 31 41 51
MDWGLHTFFI GGVNKHSTSI GKVMITVIFI FRVMILVVAA QEVWGDEQED FVNTLQPGC 60
KNVCVDHFFP VSHIRLWALQ LIFVSTPALL VAMHVAYYRH ETTRKFRRGE KRNDKDIED 120
IKKKHVRIEG SLWWTYTTSS: PFRIIFEAFF MVVFYFLYNG YHLPWVLKCG IDPCPNLVDC 180
30 FISRPTEKTV FTIFMISASV ICMLLNVAEL CYLLLKVCFR RSKRAQTQKN HPNHALKESK 240
QNMNELISD SQQNAITGFP S
```

Seq ID NO: 550 DNA sequence
Nucleic Acid Accession #: NM_002571.1
Coding sequence: 99..587

```
35 1 11 21 31 41 51
CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCGC AGCCATGCTG TGCTCCTGTC 60
TCACCCCTGG GCTGGCCCTG GTCTGTGGTG TCCCGGCCAT GGACATCCCC CAGACCAAGC 120
40 AGGACCTTGA GCTCCCAAAG TTGGCAGGGA CCTGGCAGTC CATGGCCATG GCGACCAACA 180
ACATCTCCCT CATGGCGACA CTGAAGGCCC CTCTGAGGGT CCACATCACC TCACTGTTGC 240
CCACCCCGGA GGACAAACCTG GAGATCGTTC TGACAGATG GGAGAACAAAC AGCTGTGTGT 300
AGAAGAAGGT CCTTGAGAGC AAGACTGGGA ATCCAAAGAA GTTCAAGATC AACTATACGG 360
45 TGGCGAACGA GGCCACGCTG CTGATAGTCT ACTACGACAA TTTCTGTGTT CTCTGCCTAC 420
AGGACACCAC CACCCCATC CAGAGCATGA TGTGCCAGTA CCTGGCCAGA GTCCGTGGTG 480
AGGACGATGA GATCATGCAG GGATTCATCA GGGCTTTTCA GCCCTGCCCC AGGCACCTAT 540
GGTACTTGCT GGACTTGAAG CAGATGGAAG AGCCGTGCGC TTTCTAGCTC ACCTCCGCTT 600
CCAGGAAGAC CAGACTCCCA CCCTCCACA CCTCCAGAGC AGTGGGACTT CCTCTGCCCC 660
50 TTTCAAAGAA TAACCACAGC TCAGAAGACC ATGACGTGGT CATCTGTGTC GCCATCCCTT 720
TCTCTGCTCA CACCTGCACC ATTGCCATGG GGAGGCTGCT CCTGGGGGGC AGAGTCTCTG 780
GCAGAGGTTA TTAATAAACCC CTTGGAGCAT G
```

Seq ID NO: 551 Protein sequence
Protein Accession #: NP_002562.1

```
55 1 11 21 31 41 51
MDIPQTKQDL ELPLKLAGTWH SMAMATNNIS LMATLKAPLR VHITSLLPPT EDNLEIVLHR 60
WENNNSCVERK VLGEKTNPNK KFKINYTVAN EATLLDIDYD NFLFLCLQDT TTFIQSMMQ 120
60 YLARVLVEDD EIMQGFIRAF RPLPRHLWYL LDLKQMEPC RF
```

Seq ID NO: 552 DNA sequence
Nucleic Acid Accession #: NM_006500.1
Coding sequence: 27..1967

```
65 1 11 21 31 41 51
ACTTGCCTCT CGCCCTCCGG CCAAGCATGG GGCTTCCAGG GCTGGTCTGC GCCTTCTTGC 60
TGCCTGCTGT CTGCTGCTGT CCTCGGCTCG CGGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120
70 CGCCTGAGCT GGTGGAGGTG GAAGTGGGCA GCACAGCCCT TCTGAAGTGC GGCTCTTCCC 180
AGTCCCAAGG CAACCTCAGC CATGTGCACT GGTTTTCTGT CCACAAGGAG AAGCGGACGC 240
TCATCTTCGG TGTGGGCCAG GGCCAGGGCC AGAGCGAACC TGGGGAGTAC GAGCAGCGGC 300
TCAGCCTCCA GACAGAGGGG GCTACTCTGG CCCTGACTCA AGTCAACCCC CAAGACGAGC 360
75 GCATCTTCTT GTGCCAGGGC AAGCGCCCTC GGTCCCAGGA GTACCGCATC CAGCTCCGGC 420
TCTACAAAGC TCCGAGGAGG CCAAAATCCC AGGTCAACCC CCTGGGCATC CCTGTGAACA 480
GTAAGGAGCC TGAGGAGGTC GCTACCTGTG TAGGGAGGAA CGGGTACCCC ATTCTCTAAG 540
TCATCTGTTA CAAGAAATGGC CGGCCTCTGA AGGAGGAGAA GAACCGGGTC CACATTCACT 600
CGTCCAGAC TGTGGAGTGC AGTGGTTTGT ACACCTTGCA GAGTATTCTG AAGGCACAGC 660
80 TGGTTAAAGA AGACAAAGAT GCCCAGTTTT ACTGTGAGCT: CAACTACCGG CTGCCCAGTG 720
GGAACCACTA GAAGGAGTCC AGGGAAGTCA CCGTCCCTGT TTTCTACCGG ACAGAAAAG 780
TGTGGCTGGA AGTGGAGGCC GTGGGAATGC TGAAGGAAGG GGACCGCGTG GAAATCAGT 840
GTTTGGCTGA TGGCAACCTT CCACCACTCT TCAGCATCAG CAAGCAGAAC CCCAGACCA 900
GGGAGGCAGA GGAAGAGACA ACCAAGCACA ACGGGGTCTT GGTGCTGAGC CTTGCCCGGA 960
85 AGGAACACAG TGGGCGCTAT GAATGTCAGG CCTGGAACTT GGACACCATG ATATCGCTGC 1020
TGAGTGAAAC ACAGGAACCTA CTGGTGAATC ATGTGTCTGA CTTCCGAGTG AGTCCCGCAG 1080
CCCTGTGAGC ACAGGAAGGC AGCAGCCTCA CCTGACCTG TGAGGCAGAG AGTAGCCAGG 1140
ACCTCGAGTT CCAAGTGGCTG AGAGAAGAGA CAGACCAAGT GCTGGAAGAG GGGCCTGTGC 1200
```

5
 10
 15
 20
 25
 30
 35
 40

TTCAGTTGCA TGAGCTGAAA CGGGAGGAGG GAGGCGGCTA TGCGTGGCTG GCGTCTGTGC 1260
 CCAGCATACC CGGCTGGAAC CGCACACAGC TGGTCAAGCT GGCCATTTT GGCCCCCTT 1320
 GGATGGCATT CAAGGAGAGG AAGGTGTGGG TGAAAGAGAA TATGGTGTG AATCTGTCTT 1380
 GTGAGCGTC AGGACACCCC CGGCCACCA TCTCTGGAA CGTCAACGGC ACGGCAAGTG 1440
 AACAGAGCA AGATCCACAG CGAGTCTGA GCACCTGAA TGTCCTGCTG ACCCCGGAGC 1500
 TGTGGAGAC AGGTGTGAA TGACGGCCT CCAACGACCT GGGCAAAAAC ACCAGCATCC 1560
 TCTCTCTGGA GCTGGTCAAT TTAACACACC TCACACAGA CTCACACACA ACCACTGGCC 1620
 TCAGCACTTC CACTGCCAGT CCTCATACCA GAGCCACAG CACCTCCACA GAGAGAAAGC 1680
 TGCCGGAGCC GGAGAGCCGG GGGTGGTCA TGGTGGCTGT GATTGTGTGC ATCTGGTCC 1740
 TGGCGGTGCT GGGCGCTGTC CTCTATTTC TCTATAAGAA GGGCAAGCTG CGGTGCAGGC 1800
 GCTCAGGAGAA GCAGAGATC ACCTGCCCC CGTCTCGTAA GACCGAATT GTAGTTGAAG 1860
 TTAAGTCAGA TAAGTCCCA GAAGAGATGG GCCTCTGCA GGGCAGCAGC GGTGACAAAG 1920
 GGGCTCCGGG AGACCAGGGA GAGAAATACA TCGATCTGAG GCATTAGCCC CGAATCACTT 1980
 CAGCTCCCTT CCCTGCTGG ACCATTCCCA GCTCCTGCT CACTCTTCT TCAGCCAAAG 2040
 CCTCCAAAGG GACTAGAGAG AAGCTCCTG CTCCCTCAC CTGCACACCC CTTTTCAGAG 2100
 GGCACCTGGG TTAGGAGCTG AGGACCTCAC TTGGCCCTGC AAGCCGCTT TCAGGGAGCA 2160
 GTCCACACCC ATCTCTCCA CGTTGAGTGA AGCTCATCCC AAGCAAGGAG CCCCAGTCTC 2220
 CGAGCCGGGT AGGAGAGTTT CTTGCAGAAC GTGTTTTTC TTTACACACA TTATGGCTGT 2280
 AAATACCTGG CTCCTGCCAG CAGCTGAGCT GGGTAGCCTC TCTGAGCTGG TTTCTGCCC 2340
 CAAAGGCTGG TCTCCACCAT CCAGGTGCAC CACTGAAGTG AGGACACACC GGAGCCAGGC 2400
 GCCTGCTCAT GTTGAAGTGC GCTGTTACA CCCGCTCCGG AGAGCACCCC AGCGGCATCC 2460
 AGAAGCAGCT GCAGTGTTC TGCCACCACC CTCTGCTCG CCTCTCAAA GTCTCTGTG 2520
 ACATTTTTTC TTTGGTCAGA AGCCAGGAAC TGGTGTCAAT CCTTAAAGA TACGTGCCGG 2580
 GGCCAGGTGT GGTGGCTCAC GCCTGTAATC CCAGCACTT GGGAGGCCGA GCGGGCCGGA 2640
 TCACAAAGTG AGGACGAGAC CATCTGGCT AACACGGTGA AACCTGTCT CTACTAAAAA 2700
 TACAAAAAAA AATTAGCTAG GCGTAGTGGT TGGCACCTAT AGTCCAGCT ACTCGGAAGG 2760
 CTGAAGCAGG AGAATGTTAT GAATCCAGGA GGTGGAGCTT GCATGAGCC GAGACCGTGC 2820
 CACTGCACCT CAGCCTGGGC AACACAGCGA GACTCCGTCT CGAGGAAAAA AAAAGAAAAA 2880
 ACGGTACCTT CGGTGAGGAG AGCTGGGCGC TGTTTTGAG TTCAGGTGAA TTAGCCTCAA 2940
 TCCCGTGTGT CATTGTCTCC CATAGCCCTC TTGATGATC ACGTAAACT GAAAGGCAGC 3000
 GGGGAGCAGA CAAGATGAG GTCTACACTG TCCTTCAATG GGATTAAGC TATGGTTATA 3060
 TTAGCACCAA ACTTCTACAA ACCAAGCTCA GGGCCCCAAC CCTAGAAGGG CCCAAATGAG 3120
 AGAATGTTAC TTAGGATGG AAAACGGGGC CTGGCTAGAG CTTCGGGTGT GTGTGTCTGT 3180
 CTGTGTGTAT GCATACATAT GTGTGTATAT ATGGTTTTGT CAGGTGTGTA AATTTGCAAA 3240
 TGTGTTCTCT TATATATGTA TGTATATATA TATATGAAA TATATATATA TATGAAAAA 3300
 AAGCTTAAT TGTCCAGAA AATCATACAT TGCTTTTTTA TTCTACATGG GTACCAACAGG 3360
 AACCTGGGGG CTTGTGAAC TACAACCAA AGGCACACAA AACCGTTTCC AGTTGGCAGC 3420
 AGAGATCAGG GGTACCTCT GCTTCTGAGC AAATGGCTCA AGCTCTACCA GAGCAGACAG 3480
 CTACCTACTT TTTAGCAGC AAAACGTCCT GTATGACGCA GCACGAAGGG CCTGCGCAGC 3540
 TTTAGCAGG AGCTATGTC CTCTCTATCG TTTCCGTCCA CTT

Seq ID NO: 553 Protein sequence
 Protein Accession #: NP_006491.1

45
 50
 55

1 11 21 31 41 51
 GLPRLVCAFL LAACCCCPRV AGVPGAEQAP APBLVEVEVG STALLKCGLS QSQGNLSHVD 60
 WFSVHEKERT LIFRVRQQQG QSEPEYEQR LSLQDRGATL ALTQVTPQDE RIFLCQGKRP 120
 RSQYRIQLR VYKAPPEPNI QVNPLGIPVN SKEPEEVATC VGRNGYPIPO VIWYKNGRPL 180
 KEENRNVHIO SSQTVESSGL YTLQSIKQAQ LKEDKDAQF YCELNYRLPS GNHMKESREV 240
 TVPVFYPTKE VHLSEVPVGM LKEGDRVEIR CLADGNPPPH FSISKQNPST REAEETND 300
 NGVLVLEPAR KEHSGRYECQ ANNLDTMISL LSEPEELLVN YVSDVRVSPA APERQEGSSL 360
 TLTCAESSQ DLEFQWLREE TDQVLERGPV LQLHDLKREA GGGYRCVASV PSIPGLNRTQ 420
 LVKLAIFGPP WMAFKERVV VKENMVLNLS CEASGHPRPT ISWNVNGTAS EQDQDPQRVL 480
 STLNVLTPE LLETGVECTA SNDLGKNTSI LFLELVNLT LTPDSNTTGT LSTSTASPH 540
 RANSTSTERK LPEPESRGVV IVAVIVCILV LAVLGAVLYF LYKKGKLPKR RSGKQETLTP 600
 PSRKTELVS VSKDKLPEEM GLLQSSGDK RAPGDQGEKY IDLRH

Seq ID NO: 554 DNA sequence
 Nucleic Acid Accession #: NM_003183.3
 Coding sequence: 165..2639

65
 70
 75
 80
 85

1 11 21 31 41 51
 TCGAGCCTGG CGGTAGAATC TTCCAGTAG GCGGCGCGGG AGGGAAGAAG GATTGAGGG 60
 GCTAGGCGGG GCGGATCCCG TCCTCCCCCG ATGTGAGCAG TTTCCGAAA CCGCGTCAGG 120
 CGAAGGCTGC CCAGAGAGGT GGAGTCGGTA GCGGGGCGGG GAACATGAGG CAGTCTCTCC 180
 TATTCTGAC CAGCGTGGTT CCTTTCGTGC TGGCGCGCGG ACCTCGGAT GACCCGGGCT 240
 TCGGCCCCCA CCAGAGACTC GAGAAGCTTG ATTCTTTGCT CTCAGACTAC GATATTCTCT 300
 CTTTATCTAA TATCCAGCAG CATTGGTAA GAAAAAGAGA TCTACAGACT TCAACACATG 360
 TAGAAACACT ACTAATCTTT TCAGCTTTGA AAAGGCATT TAAATTATAC CTGACATCAA 420
 GTACTGAACG TTTTTCACAA AATTTCAAGG TCGTGGTGGT GGATGGTAAA AACGAAAGCG 480
 AGTACACTGC AAAATGGCAG GACTTCTTCA CTGGACACGT GGTGGTGAG CTGACTCTA 540
 GGGTCTAGC CCACATAAGA GATGATGATG TTATAATCAG AATCAACACA GATGGGGCGG 600
 AATATAACT AGAGCCACTT TGGAGATTGG TTAATGATAC CAAAGACAAA AGAATGTTAG 660
 TTTATAAATC TGAGATATC AAGAATGTTT CACGTTTGCA GTCTCCAAA GTGTGTGGTT 720
 ATTTAAAGT GGATAATGAA GAGTTGCTCC CAAAGGGTT AGTAGACAGA GAACCACTG 780
 AAGAGCTTGT TCATCGAGTG AAAAGAAGAG CTCACCCAGA TCCCATGAAG AACCGTGTA 840
 AATTATGGT GGTAGCAGAT CATCGCTTCT ACAGATACAT GGGCAGAGGG GAAGAGAGTA 900
 CAACTACAAA TTACTTAATA GAGCTAATTG ACAGAGTTGA TGACATCTAT CGGAACACTT 960
 CATGGGATAA TCAGGTTTTT AAAGGCTATG GAATACAGAT AGAGCAGATT CGCATTTCTA 1020
 AGTCTCCACA AGAGGTAAAA CTGGTGAAA AGCACTACAA CATGGCAAAA AGTTACCCAA 1080
 ATGAAGAAAA GGATGCTTGG GATGTGAAGA TGTGTCTAGA GCAATTTAGC TTTGATATAG 1140
 CTGAGGAAGC ATCTAAAGTT TGCTTGGCAC ACCTTTTCAC ATACCAAGAT TTTGATATGG 1200
 GAACCTTGG ATTAGCTTAT GTTGGCTCTC CCAGAGCAAA CAGCCATGGA GGTGTTGTG 1260
 CAAAGGCTTA TTATAGCCCA GTTGGGAAGA AAAATATCTA TTGAATAGT GGTTTGAAGA 1320
 GCACAAAGAA TTATGGTAAA ACCATCCTTA CAAAGGAAGC TGACCTGGTT ACACTCATG 1380

	AAATTGGGACA	TAATTTTGGG	GCAGAACATG	ATCCGGATGG	TCTAGCAGAA	TGTGCCCGGA	1440
	ATGAGGACCA	GGGAGGGAAA	TATGTCTATG	ATCCCATAGC	TGTGAGTGGC	GATCACCAGA	1500
	ACAATAAGAT	GTTTTCACAA	TGCAGTAAAC	AATCAATCTA	TAAGACCATT	GAAAGTAAGG	1560
5	CCCAGGAGTG	TTTTCAAGAA	CGCAGCAATA	AAGTTTGTGG	GAACCTCGAGG	GTGGATGAAG	1620
	GAGAAGAGTG	TGATCTCTGG	ATCATGTATC	TGAACAACGA	CACCTGCTGC	AACAGCGACT	1680
	GCAAGTTGAA	GGAAGGTGTC	CAGTGCAGTG	ACAGGAACAG	TCCTTGCTGT	AAAACTGTCT	1740
	AGTTTGAAGC	TGCCCAAGAG	AAGTGCCAGG	AGCGGATTAA	TGCTACTTGC	AAAGGCGTGT	1800
	CCTACTGCAC	AGGTAATAGC	AGTGAGTGCC	CGCCTCCAGG	AAATGCTGAA	AATGACACTG	1860
10	TTTGCTTGGG	TCTTGSCAAG	TGTAAGGATG	GGAAATGCAT	CCCTTCTGCG	GAGAGGGGAA	1920
	AGCAGCTGGA	GTCCCTGTGCA	TGTAATGAAA	CTGACAACCTC	CTGCAAGGTG	TGCTGCAGGG	1980
	ACCTTTCTGG	CCGCTGTGTG	CCCTATGTCT	ATGCTGAAAC	AAAGAACTTA	TTTTTGAGGA	2040
	AAGGAAAGCC	CTGTACAGTA	GGATTTGTGT	ACATGAATGG	CAATGTGAG	AAACGAGTAC	2100
	AGGATGTAAT	TGAACGATTT	TGGGATTTC	TTGACCAGCT	GAGCATCAAT	ACTTTTGGA	2160
15	AGTTTGTAGC	AGACAACATC	GTGCGGTCTG	TCCTGGTCTT	CTCCTTGATA	TTTTGGATTTC	2220
	CTTTCCAGAT	TCTTGTCCAT	TGTGTGGATA	AGAAATTTGA	TAAACAGTAT	GAATCTCTGT	2280
	CTCTGTTTCA	CCCCAGTAAC	GTCGAAATGC	TGAGCAGCAT	GGATTCTGCA	TCGGTTCGCA	2340
	TTATCAAAACC	CTTTCCTGCG	CCCCAGACTC	CAGGCGCGCT	GCAGCCTGCC	CCTGTGATCC	2400
	CTTCGGCGCC	ACGAGCTCCA	AAACTGGACC	ACCAGAGAAT	GGACACCATC	CAGGAAGACC	2460
20	CCAGCAGAGA	CTCCCATATG	GACGAGGATG	GGTTTGAGAA	GGACCCCTTC	CCAAATAGCA	2520
	GCACAGCTGC	CAAGTCATTT	GAGGATCTCA	CGGACCATCC	GGTCGCCAGA	AGTGAAAAGG	2580
	CTGCTCTCTT	TAAACTGCAG	CGTCAGAAATC	GTGTTAACAG	CAAAGAAACA	GAGTGCTAAT	2640
	TTAGTCTCTA	GCTCTCTGTA	CTTAAGTGTG	CAAAATATTT	TTATAGATTT	GACCTACAAA	2700
	TCAATCAGAG	CTTGATATTT	GTGAAGACTG	GGAAGTGACT	TAGCAGATGC	TGGTCATGTG	2760
25	TTTGAACCTC	CTGCAGGTAA	ACAGTCTCTG	TGTGGTTTGG	CCCTTCTCTT	TTTGAAAAGG	2820
	TAAGGTGAAA	GTGAATCTAC	TTATTTTGAG	GCTTTCAGGT	TTTAGTTTCT	AAAAATCTCT	2880
	TGACCTGTG	GTGCAAAAGC	AGAAAATACA	GCTGGATTGG	GTTATGAATA	TTTACGTTT	2940
	TGTAATTAAT	TCTTTTATAT	TGATAACAGC	ACTGACTAGG	GAAATGATCA	GTTTTTTTTT	3000
	ATACACTGTA	ATGAACCGCT	GAATATGAAG	CATTGGGCAT	TTATTTGTGA	GAAAAGTGG	3060
30	ATAGTTTTTT	TTTTTTTTTT	TTTTTTTTTC	CTTCAACTAA	AAACAAGGGA	GATAAATTTA	3120
	GTATACATTG	TATCTAAATT	GTGGGTCTAT	TTCTAGTTAT	TACCCAGAGT	TTTTATGTAG	3180
	CAGGGAATAAT	ATATATCTAA	ATTTAGAAAT	CATTGGGTTT	AATATGGCTC	TTTCAATTTT	3240
	TAAGACTAAT	GCTCAGAAC	TAACCACTAC	CTTACAGTGA	GGCTATATCA	TGGTAGCCAG	3300
	TTGAATTTAT	GAAATCTACC	AACCTGTTAG	GGCCTGATT	TGCTGGGCAG	TTTTTCTGTA	3360
35	TTTTATAAGT	ATCTTCATGT	ATCCCTGTTA	CTGATAGGGA	TACATGTCTT	AGAAAATTCA	3420
	CTATTGGCTG	GGAGTGGTGG	CTCATGCCTG	TAATCCACGC	ACTTGGAGAG	GCTGAGGTTG	3480
	CGCCACTACA	CTCCAGCCTG	GGTGACAGAG	TGAGATCTGC	CTC		

Seq ID NO: 555 Protein sequence
Protein Accession #: NP_003174.2

40	1	11	21	31	41	51	
	MRQSLFLTS	VVPFVLAPRP	PDDPGFGPHQ	RLEKLDLSLS	DYDILSLSN	OQHSVRKRDL	60
45	QTSHTVETLL	TPSALKRHFH	LYLTSSTERF	SNFKVNVVD	GNSESEYAK	WQDFPTGHVV	120
	GEPDSRVLAH	IRDDVVIIRI	NTDGAENYIE	PLWRVNDTK	DKRMLVYKSE	DIKNVSRLOS	180
	PKVCGYLVVD	NEELIPKGLV	DREPPPELVH	RVKRRADPD	MKNTCKLLVV	ADHRFYRYMG	240
	RGEESTTINY	LIELIDRVDD	IYRNTSWDNA	GFKGYGQIE	QIRILKSPQE	VKPGKHYNM	300
	AKSYNPEEKD	ANDVLMLEQ	PSFDIAEAS	KVCLAHLPY	QDFDMGTGL	AVVGSPPRNS	360
50	HGGVCPKAYY	SPVGKKNLYL	NSGLTSTKNY	GKTLILTKEAD	LVTHELHGN	FGAEHDPDGL	420
	AECAPNEQED	GKYVMYPIAV	SGDHENKMF	SNCSKQSIYK	TIESKAQECF	QERSNKKVGN	480
	SRVDEGEED	PGIMVLANND	CCNSDCTLKE	GVQCSDRNSP	CCKNQCFETA	QKKCQEAANA	540
	TCKGVSYCTG	NSSECPPPGN	AENDTVCLDL	GKCKDGKCIPI	FCERBQQLS	CACNETDNCS	600
	KVCCRDLSGR	CVPIYVDAEQ	NLFRLKRGKPC	TVGFCMMNGK	CEKRVQDVIE	RFWDFIDQLS	660
55	INTFGKFLAD	NIVGVSILVPS	LIFWIPFSIL	VHCVDKLLDK	QVESLSLFHP	SNVEMLSMD	720
	SASVRIIKPF	PAPQTPGRGL	PAPVIPSAPA	APKLDHQRMD	TIQEDPSTDS	HMDDEGFED	780
	PFPSNSTAAK	SPEDLTDRHP	ARSEKAASF	LQRQNRVNSK	ETEC		

Seq ID NO: 556 DNA sequence
Nucleic Acid Accession #: NM_021832.1
Coding sequence: 164..2248

60	1	11	21	31	41	51	
	TGAGAGCTGG	CGGTAGAATC	TTCCCACTAG	GCGGCGCGGG	AGGAAAAGAG	GATTGAGGGG	60
65	CTAGGCGCGG	CGGATCCCGT	CCTCCCGCGA	TGTGAGCAGT	TTTCCGAAAC	CCGTCAGGC	120
	GAAGGCTGCG	CAGAGAGGTG	GAGTCGGTAG	CGGGGCGCGG	AACATGAGGC	AGTCTCTCCT	180
	ATTCTCTACC	AGCGTGGTTC	CTTTCGTGCT	GGCGCGCGGA	CCTCCGGATG	ACCCGGGCTT	240
	CGGCCCCCAG	CAGAGACTCG	AGAAGCTTGA	TTCTTTGCTC	TCAGACTACG	ATATTCTCTC	300
70	TTTATCTAAT	ATCCAGCAGC	ATTCCGTAAG	AAAAGAGAT	CTACAGACTT	CAACACATGT	360
	AGAAACACTA	CTAATCTTTT	CAGCTTTGAA	AAGGCATTTT	AAATTATACC	TGACATCAAG	420
	TACTGAACGT	TTTTACAAAA	ATTTCAGGT	CGTGGTGGTG	GATGGTAAAA	ACGAAAGCGA	480
	GTACACTGTG	AAATGGCAGG	ACTTCTTCAC	TGGACACGTG	GTGGTGGAGC	CTGACTCTAG	540
	GGTTCTAGCC	CACATAAGAG	ATGATGATGT	TATAATCAGA	ATCAACACAG	ATGGGCGCGA	600
75	ATATAACATA	GAGCCACTTT	GGAGATTGTG	TAATGATACC	AAAGACAAAA	GAATGTTAGT	660
	TTATAAATCT	GAGATATATCA	AGAATGTTTC	ACGTTTGCAG	TCTCCAAAAA	TGTGTGGTTA	720
	TTTAAAGTGT	GATAATGAAG	AGTTGCTCCC	AAAAGGGTTA	GTAGACAGAG	AAACCACCTGA	780
	AGAGCTGTGT	CATCGAGTGA	AAAGAAGAGC	TGACCCAGAT	CCCATGAAGA	ACACGTGTAA	840
	ATTATTGGTG	GTAGCAGATC	ATCGCTTCTA	CAGATACATG	GGCAGAGGGG	AAGAGAGTAC	900
80	AACTACAAAT	TACTTAATAG	AGCTAATTGA	CAGAGTTGAT	GACATCTATC	GGAAACACTTC	960
	ATGGGATAAT	GCAGGTTTTT	AAGGCTATGG	AATACAGATA	GAGCAGATTC	CAATTCTCAA	1020
	GTCTCCACAA	GAGGTAAGAA	CTGGTGAAAA	GCACACAAAC	ATGGCAAAAA	GTTACCCAAA	1080
	TGAAGAAAG	GATGCTGGGG	ATGTGAAGAT	GTTGCTAGAG	CAATTAGACT	TTGATATAGC	1140
	TGAGGAAGCA	TCTAAAGTTT	GCTTGGCACA	CCTTTTCA	TACCAAGATT	TTGATATGGG	1200
85	AACTCTTGGG	TTAGCTTATG	TTGGCTCTCC	CAGAGCAAA	AGCATGGAG	GTGTTTGTCC	1260
	AAAGGCTTAT	TATAGCCCAAG	TTGGGAAGAA	AAATATCTAT	TTGAATAGTG	GTTTGAACAG	1320
	CACAAGAAAT	TATGGTAAAA	CCATCTTAC	AAAGGAAGCT	GACCTGGTTA	CAACTCATGA	1380
	ATTGGGACAT	AAATTTGGAG	CAGAACATGA	TCCGATGGT	CTAGCAGAA	GTGCCCCGAA	1440

Seq ID NO: 557 Protein sequence
Protein Accession #: NP_068604.1

55 Seq ID NO: 558 DNA sequence
Nucleic Acid Accession #: NM_004994.1
Coding sequence: 20..2143

397

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

CCTTATCGCC GACAAGTGGC CCGCGCTGCC CCGCAAGCTG GACTCGSTCT TTGAGGAGCC - 1740
GCTCTCCAGG AAGCTTTTCT TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGCGTC 1800
GGTGTGGGTC CCGAGGCGTC TGGACAGGCT GGGCGTGGGA GCGGAGGTGG CCCAGGTGAC 1860
CGGGGCGCTC CGGAGTGGCA GGGGGAAGAT GCTGCTGTTC AGCGGGCGGC GCCTCTGGAG 1920
GTTGACAGCT AAGGCGCAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGGATGTT 1980
CCCGGCGGTG CCTTTGGACA CGCACGACGT CTTCCAGTAC CGAGAGAAAG CCTATTTCTG 2040
CCAGGACCGC TTCTACTGGC GCGTGAAGTC CCGGAGTGAG TTGAACCAGG TGGACCAAGT 2100
GGGCTACGTG ACCTATGACA TCCTGCAAGT CCCTGAGGAC TAGGGCTCCC GTCTCTGCTT 2160
GCAGTCCCAT GTAAATCCCC ACTGGGACCA ACCCTGGGGA AGGAGCCAGT TTGCGGATA 2220
CAAACCTGTA TTCTGTCTG GAGGAAGGGG AGGAGTGGAG GTGGGCTGGG CCCTCTCTTC 2280
TCACCTTTGT TTTTGTGTG AGTGTCTCTA ATAACTTGG ATTCTCTAAC CTTT

Seq ID NO: 559 Protein sequence
Protein Accession #: NP_004985.1

1 11 21 31 41 51
| | | | |
MSLWQLVLV LVLVGLCCFAA PRQRQSTLVL FPGDLRTNLT DRQLAEELYLY RYGYTRVAEM 60
RGESKSLGPA LLLQLQLSL PETGELDSAT LKAMRTPRCG VPDLRGFTQF EGDLEKWHEN 120
ITWIONYSE DLPRVIDDA FARAFALMSA VTPLTPTRVY SRDADIVIQF GVAEHGDGYP 180
FDGKDLGAAH APPPGPIQIG DAHFDDELW SLGKGVVPT RFGNADGAAC HFPFIFEGRS 240
YSACTTDRGS DGLPWCSTTA NYDTDRFGF CPSELYTRD GNADGKPCQF PFIFQGSYS 300
ACTTDRSDG YRNCATTANY DRDKLFGFCP TRADSTVMGG NSAGELCVFP FTFLGKEYST 360
CTSEGRDGR LWCAATSNFD SDKKWGFPCD QGYSFLVAA HEFGHALGLD HSSVPEALMY 420
PMYRFTGEPF LHKDDVNGIR HLYGPRPEPE PRPPTTTTPQ PTAFTVCPPT GPPTVHPSPR 480
PTAGPTGPPS AGPTGPPTAG PSTATTVPLS PVDDACNVNI FDAIAEIGNQ LYLFGDKGW 540
RFSEGRSRP QGPFLIADKW PALPRKLDV FEEPLSKKLF FFSGRQVWVY TGASVLGPRR 600
LDKLGADV AQVTGALRSR RGKMLFSGR RLWRFDVKAQ MVDPRSASEV DRMFPGVPLD 660
THDVFYREK AYFCQDRFYV RVSSRSBLNQ VDQVGYVTGD ILQCPED

Seq ID NO: 560 DNA sequence
Nucleic Acid Accession #: NM_000213.1
Coding sequence: 127..5385

35 40 45 50 55 60 65 70 75 80 85

1 11 21 31 41 51
| | | | |
CGCCCGCGCG CTGCAGCCCC ATCTCCTAGC GGCAGCCAG GCGCGGAGGG AGCGAGTCCG 60
CCCGAGGTA GGTCCAGGAC GGGCGCACAG CAGCAGCCGA GGCTGGCCGG GAGAGGGAGG 120
AAGAGGATGG CAGGGCCACG CCCCAGCCCA TGGGCCAGGC TGCTCCTGGC AGCCTTGATC 180
AGCGTCAGCC TCTCTGGGAC CTTGGCAAAC CGCTGCAAGA AGGCCCCAGT GAAGAGCTGC 240
ACGGAGTGTG TCCGTGTGGA TAAGGACTGC GCCTACTGCA CAGACGAGAT GTTCAGGGAC 300
CGGCGCTGCA ACACCCAGGC GGAGCTGCTG GCCCGGGGCT GCCAGCGGGA GAGCATCTGT 360
GTCTGAGAGA GCAGCTTCCA AATCAGAGAG GAGACCCAGA TTGACACCAC CCTGCGGCGC 420
AGCCAGATGT CCCCCAAGG CTTGCGGGTC CGCTCTGGGC CCGGTGAGGA GCGGCATTTT 480
GAGCTGGAGG TGTTTGAGCC ACTGGAGAGC CCGGTGGACC TGATACATCT CATGGAATTC 540
TCCAACCTCA TGTCCGATGA TCTGGACAAC CTTCAAGAAG TGGGGCAGAA CCTGGCTCGG 600
GCTCTGAGCC AGCTCACCAG CGACTACACT ATTGGATTGG GCAAGTTTGT GGACAAAGTC 660
AGCGTCCGCG AGACGGACAT GAGGCTCTGAG AAGCTGAAGG AGCCCTGGCC CAACAGTGAC 720
CCCCCTTCT CTTCAAGAA CGTCATCAGC CTGACAGAAAG ATGTGGATGA GTTCCGGAAT 780
AAATGCAAGG GAGAGCGGAT CTGAGGCAAC CTGGATGCTC CTGAGGGCGG CTTCGATGCC 840
ATCCTGCAGA CAGCTGTGTG CAGAGGGAC ATTGGCTGGC GCCCGGACAG CACCCACCTG 900
CTGGTCTTCT CCACCGAGTC AGCCTTCCAC TATGAGGCTG ATGGGCCCAA CGTGTCTGCT 960
GGCATCATGA CGCGCAACGA TGAACGGTGC CACCTGGACA CCACGGGCAC CTACACCCAG 1020
TACAGGTCAC AGGATACCC GTGGGTGCCC ACCCTGGTGC GCCTGTCTGC CAAGCACAAC 1080
ATCATCCCCA TCTTTGCTGT CACCAACTAC TCCTATAGCT ACTACAGGAA GCTTCACACC 1140
TATTTCCCTG TCTCTCACT GGGGGTGTG CAGGAGSACT CGTCCAACT CGTGGAGCTG 1200
CTGAGGAGG CCTTCAATCG GATCCGCTCC AACCTGGACA TCGGGGCCCT AGACAGCCCC 1260
CGAGGCGCTT GGCAGAGGT CACCTCCAAG ATGTTCCAGA AGACGAGGAC TGGGTCTTTT 1320
CACATCCGGC GGGGGGAAGT GGGTATATAC CAGGTGCAGC TGGGGGCCCT TGAGCAGCTG 1380
GATGGGAGCG ACCTGTGCCA GCTGCCGAG GACCAGAGG GCAACATCCA TCTGAAACCT 1440
TCCTTCTCCG ACGGCCTCAA GATGGACGCG GGCATCATCT GTGATGTGTG CACCTGCGAG 1500
CTGCAAAAG AGGTGCGTTC AGCTCGCTGC AGCTTCAAGG GAGACTTCGT GTGCGGACAG 1560
TGTGTGTGCA GCGAGGGCTG GAGTGGCCAG ACCTGCAACT GCTCCACCGG CTCTCTGAGT 1620
GACATTACAG CTTGCTCTCG GGAGGGCGAG GACAAGCCGT GCTCCGCGCG TGGGGAGTGC 1680
CAGTGGGCGC ACTGTGTGTG CTACGGCGAA GCGCGCTAGG AGGGTCAGTT CTGCGAGTAT 1740
GACAACTTCC AGTGTCCCGG CACTTCCGGG TTCCTCTGCA ATGACCGAGG ACCTGCTCC 1800
ATGGGCGAGT GTGTGTGTGA GCCTGGTTGG ACAGGCCCAA GCTGTGACTG TCCCTCAGC 1860
AATGCCACCT GCATCGACAG CAATGGGGGC ATCTGTAATG GAGTGGCCA CTGTGAGTGT 1920
GGCCGCTACG ACTGCCACCA GCAGTCTCTC TACAGGACA CCATCTGCGA GATCAACTAC 1980
TCGGGATACC ACCCGGGCCT CTGCGAGGAC CTACGCTCCT GCGTGCAGTG CCAGGCGTGG 2040
GGCACCAGCG AGAAGAAGGG GCGCACTGTG GAGGAATGCA ACTTCAAGGT CAAGATGTGT 2100
GACGAGCTTA AGAGAGCCGA GGAGGTGGTG GTGCGCTGCT CTTCCGCGGA CGAGGATGAC 2160
GACTGCACCT ACAGTACAC CATGGAAGGT GACGGCGCCC CTGGGCCCAA CAGCACTGTC 2220
CTGGTGACA AGAAGAAGCA CTGCCCTCCG GGCTCTTCT GGTGGCTCAT CCCCCTGCTC 2280
CTCTCTCTCC TGCGCTCTCT GCGCTGTCTA CTGCTGCTAT GCTGGAAGTA CTGTGCTGTG 2340
TGCAAGGCTT GCCTGGCACT TCTCCGCTGC TGAACCCGAG GTCACATGGT GGGCTTTAAG 2400
GAAGACACT ACATGCTGCG GGAGAACCTG ATGGCTCTG ACCACTTGGG CAGGCCCATG 2460
CTGCGCAGCG GGAACTCAA GGGCGGTGAC GTGGTCCGCT GGAAGGTGAC CAACAACATG 2520
CAGCGGCTCG GCTTTGCCAC TCATGCCGCC AGCATCAACC CCACAGAGCT GGTGGCCCTAC 2580
GGGCTGTCTT TGCGCTTGGC CGGCTTTTGC ACCGAGAACC TGCTGAAGCC TGACACTCGG 2640
GAGTGGCGCC AGCTGCGCCA GGAGGTGGAG GAGAACCCTGA ACGAGGTCTG CAGGCAGATC 2700
TCCGTGTGAC ACAAGCTCCA GCAGACCAAG TTCCGGCAGC AGCCCAATGC CGGAAAAAAG 2760
CAAGACCACT CATTTGTGGA CACAGTGTCT ATGGCGCCCC GCTCGGCCAA CGCGGCCCTG 2820
CTGAAGCTTA CAGAGAACTA GGTGGAACAG AGGGCTTTC ACAGCTCAA GGTGGCCCTC 2880
GGCTACTACA CCTTCACTGC AGACAGGAC GCGCGGGGCA TGTGGAGTGT CCAGGAGGGC 2940
GTGGAGCTGG TGGAGTACG GGTGCCCTTC TTTATCCGCG CTGAGGATGA CAGCAGAAAG 3000
CAGCTGCTGG TGGAGCCAT CGACGTGCCC GCAGGCACTG CACCTCTCGG CCGCGGCTG 3060

5
10
15
20
25
30
35
40
45

Seq ID NO: 561 Protein sequence
Protein Accession #: NP_000204.1

1 11 21 31 41 51

MAGPRPSPWA RLLLAALISV SLSGTLANRC KKAPVKSCTE CVRVKDCAY CTDEMFRDRR 60
CMTQABLLAA GCQRESIVVM ESSFQITEET QIDTTLRRSQ MSPQGLRVRL RPEGERHFEL 120
EVFEPLESVP DLVILMDFSN SMSDDLNLK KMGNLQARVL SOLTSDYITG FGFVFDKVSFV 180
PQTDMPERKL KEPWPNSDPP FSPKNVISLT EDVDEFKNKL QGERISGNLD APEGGFDAIL 240
QTAVCTRDIG WRPDSTHLIV FSTESAFHYE ADGANVLAGE MSRNDERCHL DTTGTYTQYR 300
TQDYPSPVPTL VRLLAGHNII PIFAVTNYSY SYEKLHTYF PVSSGLVLQE DSSNIVELLE 360
EAFNRIRSNL DIRALDSPRG LRTEVTSKMP KQRTTGSFHI RRGVGIYQV QLRALHVDG 420
THVCQLPEDQ KGNHILKPSF SDGLKMDAGI ICDVCTCELQ KEVRSARCSF NGDFVOGQCV 480
CSEGSQGTTC NCSTGSLSDI PQCLREGEDK PCSGRGECQC GHCVCYGEGR YEGQFCEYDN 540
FQCPRTSGFL CNDRGRCSMG QCVCEPWTG PSCDCPLSNA TCIDSNGGIC NGRGHCBCCR 600
CHCHQQLSYT DTICEINYSY IHPGLCEDLR SCVQCQAGT GEKKGRTECE CNFKVKMVDL 660
LKRAEIVVVR CSFRDEDDC TYSYTMEDGD APGPNSTVLV HKKKDCPPGS FWLILPLLLL 720
LLPLALLLLL LCKKYCACCK ACLALLPCCN RGHMVGFKED HYMLRENMA SDHLDTPLML 780
SGNLKGRDVV RWKVTNNMQR PGFATHAASI NPTELVPYGL SLRLARLCTE NLLKPDTRC 840
AQLRQEVVEN LNEVYRQISG VHKLQQTFR QPNAGKKQD HTIVDTVIMA PRSAKPALLK 900
LVEAIDVPAG TATLGRRLVN ITIIEQARD VVSFEQPEFS VSRGDQVARI FVIRRVLDGG 1020
KSQVSYRTQD GTAQGNRDYI PVEGELLFQP GEAWKELQVK LLELQVDSL LRGRQVRPFH 1080
VQLSNPKFGA HLGQPHSTTI IIRDPDELDR SFTSQMLSSQ PPHGDLGAP QNPNAKAAGS 1140
RKHFHNLWLF SGKPMGRVVK YWIQDSESE AHLSDSKVPS VELTNLYPYC DYEMKVCAYG 1200
AQEGEPYSSL VSCRTHQEVV SEPGRLAFNV VSSVTQLSW AEPAETNGEI TAYEVCYGLV 1260
NDDNRPIGPM KKVLDVNPKN RMLLIENLRE SQPYRYTVKA RRGAGWGPER EAIINLATQP 1320
KRPMSPPIIP DIPIVDAQSG EDYDSFLMYS DDVLRSPSGS QRPVSDDTE HLVNGRMDFA 1380
FPGSTNSLHR MITTSAAAYG THLSPHVPHR VLSTSSTLTR DYNLSLRSEH SHSTTLPRDY 1440
STLTSVSSH D SRLTAGVPDT PTRLVFSALG PTLRLVSWQE PRCEPLQGY SVEYQLLNGG 1500
ELHRLMIPNP AQTSVVVVDL LPNHSYVFRV RAQSQEGWGR EREGVITIES QVHPQSPLCF 1560
LPGSAFTLST PSAPGPLVFT ALSPDSLQLS WERPRFPNGD IVGLVTCME AQGGGPATAF 1620
RVDGDSPESS LTVPLGSENV PYKFKVQART TEGFGPEREG IITIESQDGG PFPQLGSRAG 1680
LFQHPQLQSEY SSITTHHTSA TEPFLVDGPT LGAQHLEAGG SLTRHVTQEF VSRTLTSSTG 1740
LSTHMDQQFF QT

Seq ID NO: 562 DNA sequence
Nucleic Acid Accession #: NM_013332.1
Coding sequence: 1..63

85

1 11 21 31 41 51

GCACGAGGGC GCTTTTGTCT CCGGTGAGTT TTGTGGCGGG AAGCTTCTGC GCTGGTGCTT 60
AGTAACCGAC TTTCCTCCGG ACTCCTGCAC GACCTGCTCC TACAGCCGGC GATCCACTCC 120
CGGCTGTTCC CCGGAGGGGT CCAGAGGCCT TTCAAGAGGA GAAGGCAGCT CTGTTTCTCT 180
GCAGAGGAGT AGGGTCTCTT CAGCCATGAA GCATGTGTTG AACCTCTACC TGTTAGGTGT 240
GGTACTGACC CTACTCTCCA TCTTGGTATG AGTGATGGAG TCCCTAGAAG GCTTACTAGA 300
GAGCCCATCG CCTGGGACCT CCTGGACCAC CAGAAGCCAA CTAGCCAACA CAGAGCCAC 360
CAAGGCCTT CCAGACCATC CATCCAGAAG CATGTGATAA GACCTCTCTC CATACTGGCC 420
ATATTTTGA ACATGACCT AGACATGTCC AGATGGGAGT CCCATTCTTA GCAGACAAGC 480
TGAGCCCGT TGAACCCAGA GAACATATTAC TAGGCCTTGA AGAACCTGTC TAACCTGGATG 540
CTCATTTGCC GGGCAAGGCC TGTTAGGCC GGTGCGGTG GCTCATGCCT GTAATCTTAG 600
CACTTTGGGA GGCTGAGGTG GGTGGATCAC CTGAGGTCAG GAGTTCGAGA CCAGCCTCGC 660
CAACATGGCG AAACCCCATC TCTACTAAAA ATACAAAAGT TAGCTGGGTG TGGTGGCAGA 720
GGCCTGTAAT CCCAGTTCTT TGGGAGGCTG AGGCGGGAGA ATTGCTTGAA CCGCGGGACG 780
GAGGTTGCAG TGAACCCAGA TCGCACTGCT GTACCCAGCC TGGGCCACAG TGCAAGACTC 840
CATCTCAAAA AAAAAAGAA AAGAAAAAGC CTGTTTAATG CACAGGTGTG AGTGGATTGC 900
TTATGGCTAT GAGATAGGTT GATCTCGCCC TTACCCCGGG GTCTGGTSTA TGCTGTGCTT 960
TCCTCAGCAG TATGGCTCTG ACATCTCTTA GATGTCCCAA CTTCAGCTGT TGGGAGATGG 1020
TGATATTTTC AACCTTACTT CCTAAACATC TGCTGGGGT TCCTTTAGTC TTGAATGTCT 1080
TATGCTCAAT TATTGGTGTG TGAGCCTCTC TTCCACAAGA GCTCCTCCAT GTTTGGATAG 1140
CAGTTGAAGA GGTGTGTGGG GTGGGCTGTT GGGAGTGAGG ATGGAATGTT CAGTGCCCAT 1200
TTCTCATTTT ACATTTTAAA GTCTTCTCTC CAACATAGTG TGTATTGGTC TGAAGGGGGT 1260
GGTGGATGC CAAAGCCTGC TCAAGTTATG GACATTGTGG CCACCATGTG GCTTAAATGA 1320
TTTTTCTAA CTAATAAAGT GGAATATATA TTTCAAAAA AAAAAAAAAA AA

Seq ID NO: 563 Protein sequence
Protein Accession #: NP_037464.1

1 11 21 31 41 51
MKHVLNLYLL GVVLTLISIF VRVMESLEGL LESPSPGTGW TTRSQLANTE PTKGLPDHPS 60
RSM

Seq ID NO: 564 DNA sequence
Nucleic Acid Accession #: NM_023915.1
Coding sequence: 250..1326

1 11 21 31 41 51
GGCAGGAGGG TTTCGTTTTC ATGCTTTACC AGAAAAATCCA CTTCCTGCC GACCTTAGTT 60
TCAAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG 120
GTGAATGGAC AGCCAGCCAC CACAATGAAA GAAATCAAAC CAGGAATAAC CTATGCTGAA 180
CCCAAGCCTC AATCGTCCCC AAGTGTTCCT TGACACGCAT CTTTGCTTAC AGTGCAATCAC 240
AAGTGAAGAA TGGGGTTCAA CTGACGCTT GCAAAATATC CAAATAACGA GCTGCACGGC 300
CAAGAGAGTC ACAATTCAGG CAACAGGAGC GACGGGCCAG GAAAGAACAC CACCCTTCAC 360
AATGAATTG ACACAATTGT CTGCGCGGTG CTTTATCTCA TTATATTGT GGCAAGCATC 420
TTGCTGAATG GTTTAGCAGT GTGGATCTTC TTCCACATTA GGAATAAAAC CAGCTTCATA 480
TTCTATCTCA AAAACATAGT GGTGTCAGAC CTCATAATGA CGCTGACATT TCCATTTOGA 540
ATAGTCCATG ATGCAGGATT TGGACCTTGG TACTTCAAGT TTATTCTCTG CAGATACACT 600
TCAGTTTGT TTTATGCAAA CATGTATACT TCCATCGTGT TCCTTGGGCT GATAAGCATT 660
GATCGCTATC TGAAGGTGGT CAAGCCATTT GGGGACTCTC GGATGTACAG CATAACCTTC 720
ACGAAGGTTT TATCTGTTTG TGTGTTGGTG ATCATGGCTG TTTTGTCTTT GCCAAACATC 780
ATCCTGACAA ATGGTCAGCC AACAGAGGAC AATATCCATG ACTGCTCAA ACTTAAAGT 840
CCTTTGGGGG TCAAATGGCA TACGGCAGTC ACCTATGTGA ACAGCTGCTT GTTTGTGGCC 900
GTGCTGGTGA TCTGATCGG ATGTTACATA GCCATATCCA GGTACATCCA CAAATCCAGC 960
AGGCAATTCA TAAGTCAGTC AAGCCGAAAG CGAAACATA ACCAGAGCAT CAGGGTTGTT 1020
GTGGCTGTGT TTTTACCTT CTTTCTACCA TATCACTGTG GCAGAATTCC TTTTACTTTT 1080
AGTCACCTAG ACAGGCTTTT AGATGAATCT GCACAAAAA TCCTATATTA CTGCAAGAA 1140
ATTACACTTT TCTTGTCTGC GTGTAATGTT TGCTGGATC CAATAAITTA CTTTTCATG 1200
TGTAGGTCAT TTTCAAGAAG GCTGTTCAAA AATCAAATA TCAGAACGAG GAGTGAAAGC 1260
ATCAGATCAC TGCAAGTGT GAGAAGATCG GAAGTTCGCA TATATTATGA TTACACTGAT 1320
GTGTAGGCCT TTTATTGTTT GTTGAATCG ATATGTACAA AGTGTAAATA AATGTTTCTT 1380
TTCATTATCC TTAATAAAAA AA

Seq ID NO: 565 Protein sequence
Protein Accession #: NP_076404

1 11 21 31 41 51
MGFNLTLAKL PNNELHGES HNSGNRSDGP GRNTTLRNEF DTIVLPVLYL IIFVASILLN 60
GLAVNIFFHI RNKTSFIFYL KNIIVADLIM TLTFFPRIVH DAGFGPWYFK FILCRYTSVL 120
FYANMYTSIV FLGLISIDRY LKVVKPFQDS RMYSTFTKV LSVCVWVIMA VLSLPNIIIT 180
NQPTEDNIH DCSKLKSPLG VKWHTAVTVV NSCLFVAVLV ILIGCYIAIS RYIHKSSRF 240
ISQSSRKRIE NQSIKRVVAV FFTCFLPYHL CRIPFTPSHL DRLLDESAQK ILIYCKEITL 300
FLSACNVCLD PIIFYFFMCRS PSRRLFKKSNI IRTSESIRS LQSVRRSEVR IYYDYTDV

Seq ID NO: 566 DNA sequence
Nucleic Acid Accession #: NM_005365.1
Coding sequence: 1..948

1 11 21 31 41 51
ATGCTCTCTG AGCAGAGSAG TCCGCACTGC AAGCCTGATG AAGACCTTGA AGCCCAAGGA 60
GAGGACTTGG GCGCTGATGG TGACAGAGAA CCCACAGGCG AGGAGGAGGA GACTACCTCC 120
TCCTCTGACA GCAAGGAGGA GGAGGTGTCT GCTGCTGGGT CATCAAGTCC TCCCAGAGT 180
CCTCAGGAG GCGCTTCTCT CTCCATTTC GTCTACTACA CTTTATGGAG CCAATTGAT 240
GAGGGCTCCA GCAGTCAAGA AGAGGAAGAG CCAAGCTCCT CGTCCGACC AGCTCAGCTG 300
GAGTTCATGT TCCAAGAAGC ACTGAAATTG AAGTGGCTG AGTTGGTICA TTTCTGCTC 360

5
10
15
20
25
30
35
40
45
50
55
60

CACAAATATC GAGTCAAGGA GCGGTCACA AAGGCAGAAA TGCTGGAGAG CGTCATCAA 420
AATTACAAGC GCTACTTTCC TGATGATCTT GGCAGAGCCT CCGAGTTCAT GCAGGTGATC 480
TTTGGCACTG ATGTGAAGGA GGTGGACCCC GCGGCGCACT CCTACATCCT TGTCAGTGCT 540
CTTGGCCTCT CGTGGCAGAT CATGCTGGGT GATGGTCATA GCATGCCCAA GGCGCCCTC 600
CTGATCATTG TCTGGGTGT GATCCTAACC AAAGACAAC TGCCTCTGA AGAGGTATC 660
TGGGAAGCGT TGAGTGTGAT GGGGGTGTAT GTTGGGAAG AGCACATGTT CTACGGGAG 720
CCGAGGAAGC TGCTACCCA AGATTGGGTG CAGGAAAAC ACCTGGAGTA CGGCAGGTG 780
CCGCGCAGTG ATCCTGGCA CTAAGAGTTC CTGTGGGTG CCAAGGCCA CGCTGAAACC 840
AGCTATGAG AGGTATATA TTATTGGTC ATGCTCAATG CAAGAGAGCC CATCTGCTAC 900
CCATCCCTTT ATGAAGAGGT TTTGGGAGAG GAGCAAGAG GAGTCTGA

Seq ID NO: 567 Protein sequence
Protein Accession #: NP_005356.1

15
20
25
30
35
40
45
50
55
60

1 11 21 31 41 51
MSLEQRSPHC KPDEDLBAQG EDLGLMGAQE PTGEEETTS SSDSKEEEVS AAGSSPPQS 60
PQGGASSIS VYVTLWSQD EGSSSQEEEE PSSVDPAQL EFMFQEAALK KVAELVHFL 120
HKYRVKEPVT KAEMLESVIK NYKRYFPVIF GKASEFMQVI FGTDVKEVDP AGHSYILVTA 180
LGLSCDSMLG DGHSMKPAAL LIIVLGVLIT KDNCAPPEVI WEALSVMGVY VGKEHMFYGE 240
PRKLLTQDWV QENYLEYRQV PGSDPAHYEF LWGSKAHAET SYEKVINYL VMLNAREPICY 300
PSLYEEVLGE EQBQV

Seq ID NO: 568 DNA sequence
Nucleic Acid Accession #: NM_014400
Coding sequence: 86..1126

15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
GGTTACTCAT CCTGGGCTCA GGTAAGAGGG CCGAGCTCG GAGGCGGCAC ACCCAGGGGG 60
GAGCCCAAGG GAGCAGGACG GAGCCATGGA CCCCGCCAGG AAAGCAGGTG CCCAGGCCAT 120
GATCTGGACT CGAGGCTGGC TGCTGCTGCT GCTGCTTCGC GGAGGAGCGC AGGCCCTGGA 180
GTGCTACAGC TGCGTGACGA AAGCAGATGA CGGATGCTCC CGAACAAGA TGAAGACAGT 240
GAAGTGGCGG CCGGCGGTGG ACCTCTGCAC CGAGGCGGTG GGGCGGTGG AGACCATCCA 300
CGGACAATTC TCGTGGGAG TGCSGGGTG CGGTTCCGGA CTCCCAGCA AGAATGACCG 360
CGGCTGGGAT CTTACGGGC TTCTGGCGTT CATCCAGCTG CAGCAATGCG CTCAGGATCG 420
CTGCAACGCC AAGCTCAACC TCACCTCGCG GCGCTCGAC CCGCAGGTA ATGAGAGTGC 480
ATACCGCGCC AACCGGTGG AGTGCTACAG CTGTGTGGGC CTGAGCCGGG AGGCGTGCCA 540
GGGTACATCG CCGCGGTGCG TGAGCTGCTA CAAGCCAGC GATCATGTCT ACAAGGGCTG 600
CTTCGACGGC AAGCTCACTT TGAAGGAGC TAATGTGACT GTGTCTTTCG CTGTCGGGG 660
CTGTGTCAG GATGAATCTT GCACTCGGGA TGGAGTAACA GGGCCAGGGT TCACGCTCAG 720
TGCTCTCTGT TGCCAGGGGT CCGCTGTAA CTCTGACCTC CGCAACAAGA CTACTTCTC 780
CCCTCGAATC CCACCGCTTG TCGGCTGCC CCCTCCAGAG CCCAGACTG TGGCTCAAC 840
CACATCTGTC ACCACTTCTA CCGCGCCCC AGTGAGACC ACATCCACCA CCAAAACCAT 900
GCCAGGCCCA ACCAGTCAGA CTCGAGACA GGGAGTAGAA CAGAGGCCT CCGGGATGA 960
GGAGCCCAAG TTGACTGGAG CCGCGCTGG CCACAGGAC CGCAGCAAT CAGGGCAGTA 1020
TCCTGCAAAA GGGGGGCCCC AGCAGCCCCA TAATAAAGGC TGTGTGGCTC CCACAGCTGG 1080
ATTGGCAGCC TTCTGTGTG CCGTGGCTGC TGGTGTCTTA CTGTGAGCTT CTCCACCTGG 1140
AAATTTCCCT CTCACCTACT TCTCTGGCCC TGGGTACCCC TCTTCTATC ACTTCTCTGT 1200
CCCACCACTG GACTGGGCTG GCCCAGCCCC TGTTTTCCA ACATTCCCA GTATCCCCAG 1260
CTTCTGCTGC GCTGTTTGG GCGTTTGGGA AATAAAATAC CGTTGTATAT ATTCTGGCAG 1320
GGGTGTTCTA GCTTTTGTAG GACAGCTCCT GTATCCTTCT CATCCTTGTC TCTCCGCTTG 1380
TCCTCTGTG ATGTAGGAC AGAGTGAGAG AAGTCAGCTG TCACGGGGAA GGTGAGAGAG 1440
AGGATGCTAA GCTTCTACT CACTTCTCC TAGCCAGCCT GGACTTTGGA CGGTGGGGTG 1500
GGTGGGACAA TGGCTCCCCA CTCTAAGCAC TGCTCTCCCT ACTCCCGCA TCTTTGGGA 1560
ATCGGTTCCC CATATGCTT CTTACTAGA CTGTGAGCTC CTCGAGGGCA GGGACCGTGC 1620
CTTATGCTG TGTGTGATCA GTTCTGGCA CATAAATGCC TCAATAAGA TTAATTACT 1680
TTGTATAGTG AAAAAA

Seq ID NO: 569 Protein sequence
Protein Accession #: NP_055215

15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
MDPARKAGAQ AMIWTAGWLL LLLLRGAQA LECYSCVQKA DGCSPNKKM TVKCAPGV 60
CTEAVGAVET IHGQFSLAVX GCGSLPGKN DRGLDLHGLL AFIQLQCAQ DRCAKLNLT 120
SRALDPAGNE SAYPPNGVEC YSCVGLSREA CQGTSPPVVS CYNASDHVYK GCFDGNVTLT 180
AANVTVSLPV RGCVDDEFCT RDGVTGPFT LSGSCCGSR CNSDLRNKTY FSPRIPLVR 240
LPPPEPTTVA STTSVTSTTS APVRPTSTTK PMPAPTSTP RQGVHEASR DEEPRLTGGA 300
AGHQDRSNG QYPAKGGPQQ PHNKGCVAPT AGLAALLAV AAGVLL

Seq ID NO: 570 DNA sequence
Nucleic Acid Accession #: NM_005329.1
Coding sequence: 1..1662

15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
ATGCCGCTGC AGCTGACGAC AGCCCTCGCT GTGGTGGGCA CCAGCCTGTT TGCCCTGGCA 60
GTGCTGGGTG GCATCTGGC AGCCTATGTG ACGGCTACC AGTTTCATCCA CACGAAAAG 120
CACTACCTGT CTTTGGGCT GTACGGGCCC ATCTGGGCC TGCACCTGCT CATTGAGAGC 180
CTTTTGGCT TCCTGGAGCA CCGGCGCATG CGAGTGCAG GCCAGGCCCT GAAGCTGCCC 240
TCCCGCGGCG GGGGCTCGGT GGCACGTGTC ATTGCCCGT ACCAGGAGGA CCTGACTAC 300
TTGCGCAAGT GCCTGCGCTC GCGCCAGCGC ATCTCTTCC CTGACCTCAA GGTGGTCATG 360
GTGGTGGATG GCAACCGCCA GGAGGACGCC TACATGCTGG ACATCTTCCA CGAGGTGCTG 420
GGCGGACCG AGCAGGCGCG CTCTTTGTG TGGCGCAGCA ACTTCCATGA GGCAGGCGAG 480
GGTGAGACCG AGGCGACGCT GCAGGAGGGC ATGGACCGTG TCGGGATGT GGTGCGGGCC 540
AGCACCTTCT CGTGATCAT GCAGAAGTGG GGAGGCAAGC GCGAGGTGAT GTACACGGCC 600

20 Seq ID NO: 571 Protein sequence
Protein Accession #: NP_005320.1

Seq ID NO.: 572 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 148-7095

402

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

AGCCTTGCTC AGTATTCTGA TGTGCTGCC ACTACTCATG CTGCTTCAGA GACGCTGGAA 2820
 TTTGGTAGTG AATCTGGTGT TCTTTATAAA AOCCTTATGT TTTCTCAAGT TGAACCAACC 2880
 AGCAGTGATG CCATGATGCA TGCACGTTCT TCAGGGGCGT AACCTTCTTA TGCCTTGTCT 2940
 GATAATGAGG GCTCCCAACA CATCTTCACT GTTCTTACA GTTCTGCAAT ACCTGTGCAT 3000
 GATTCTGTGG GTGTAACCTA TCAGGGTTC TTATTTAGCG GCCCTAGCCA TATACCAATA 3060
 CCTAAGTCTT CGTAAATAAC CCCAAGTCA TCATTACTGC AGCCTACTCA TGCCTCTCT 3120
 GGTGATGGGG AATGGTCTGG AGCCTCTTCT GATAGTGAAT TTCTTTTACC TGACACAGAT 3180
 GGGCTGACAG CCCTTAACAT TTCTTCACTT GTTCTGTAG CTGAATTTAC ATATACAACA 3240
 TCTGTGTTTG GTGATGATAA TAAGGCGCTT TCTAAAAGTG AAATAATATA TGGAAATGAG 3300
 ACTGAACTGC AAATTCCTTC TTTCAATGAG ATGCTTTACC CTCTGAAAG CACAGTCATG 3360
 CCCAATCATG ATGATAATGT AAATAAGTTG AATGCGTCTT TACAAGAAC CTCTGTTTCC 3420
 ATTTCTAGCA CCAAGGCGAT GTTCCAGGG TCCCTTGCTC ATACCAACC TAAGGTTTTT 3480
 GATCATGAGA TTAGTCAAGT TCCAGAAAAT AACTTTTCAG TTCAACCTAC ACATACGTG 3540
 TCTCAAGCAT CTGGTGACAC TTCGCTTAAA CCTGTGCTTA GTGCAAACTC AGAGCCAGCA 3600
 TCCTCTGACC CTGCTTCTAG TGAATGTTA TCTCCTTCAA CTCAGCTCTT ATTTTATGAG 3660
 ACCTCAGCTT CTTTTACTAG TGAAGTATTG CTACAACCTT CCTTTCAGGC TTCTGATGTT 3720
 GACACCTTGC TTAACACTGT TCTCCAGCT GTGCCAGTG ATCCAATAT GGTGAAACC 3780
 CCCAAGTTG ATAAATTAG TTCTACAATG TTGCATCTCA TTGTATCAAA TCTCTCTCA 3840
 AGTGAACAAC TGCTGCACTC TACATCTGTA CCAGTTTTTG ATGCTGCGCC TACTTCTCAT 3900
 ATGCACCTG CTTCACCTCA AGGTTTGACC ATTTCTATG CAAGTGAGAA ATATGAACCA 3960
 GTTTTGTAA AAAGTGAAGT TTCCCAACCA GTGGTACCTT CTTGTACAG TAATGATGAG 4020
 TTGTTCCAAA CGGCCAATTG GGAGATTAA CAGGCCCATC CCCCAAAAGG AAGGCATGTA 4080
 TTTGCTACAC CTGTTTTATC AATTGATGAA CCATTAAATA CACTAATAAA TAAGCTTATA 4140
 CATTCGATG AAATTTTAA CTTCCACCAA AGTTCTGTGA CTGTAAGGT ATTTGCTGGT 4200
 ATTTCAACAG TTGCTTCTGA TACATTGTA TCTACTGATC ATTCTGTTC TATAGGAAAT 4260
 GGGCATGTTG CCATTACAGC TGTCTCTGCC CACAGAGATG GTTCTGTAAC CTCACAAAG 4320
 TTGCTGTTTC CTCTTAAGGC AACTTCTGAG CTGAGTCATA GTGCCAATC TGATGCCGT 4380
 TTAGTGGGTG GTGGTGAGAA TGGTGACACT GATGATGATG GTGATGATGA TGATGATGAC 4440
 AGAGGTAGTG ATGGCTTATC CATTCAATA TGTATGTCAT GCTCATCTTA TAGAGAAATCA 4500
 CAGGAAAGG TAATGAATGA TTCAGACACC CACGAAACCA GTCTTATGGA TCAGAAATAT 4560
 CCAATCTCAT ACTCACTATC TGAGAATTCT GAAGAAGATA ATAGAGTCAC AAGTGTATCC 4620
 TCAGACAGTC AAATCTGGAT GGACAGAAAG CCTGGTAAAT CACCATCAGC AAATGGGCTA 4680
 TCCCAAAAGC ACAAATGATG AAAAGAGGAA AATGACATTC AGACTGGTAG TGCTCTGCTT 4740
 CCTCTCAGCC CTGATCTTAA AGCATGGGCA GTTCTGACAA GTGATGAAGA AAGTGGATCA 4800
 GGGCAAGGTA CCTCAGTAG CCTTAATGAG AATGAGACTT CCACAGATTT CAGTTTGTGA 4860
 GACACTAATG AAAAGATGTC TGATGGGATC CTGGCAGCAG GTGACTCAGA AATAACTCCT 4920
 GGATTCACAC AGTCCCAAC ATCATCTGTT ACTAGCGAGA ACTCAGAAAT GTTCCAGCTT 4980
 TCAGAGGCA AGGCCAGTAA TAGTAGCCAT GAGTCTCGTA TTGGTCTAGC TGAGGGGTTG 5040
 GAATCCGAGA AGAAGGCAGT TATACCCCTT GTGATGCTGT CAGCCCTGAC TTTTATCTGT 5100
 CTAGTGGTTC TTGTTGGTAT TCTCATCTAC TGGAGGAAAT GCTTCCAGAC TGACACATTT 5160
 TACTTAGAGG ACAGTACATC CCCTAGAGTT ATATCCACAC CTCCAACACC TATCTTTCCA 5220
 ATTTCAAGATG ATGTCGAGC AATTCCAATA AAGCACTTTC CAAAGCATGT TGCAGATTTA 5280
 CATGCAAGTA GTGGGTTTAC TGAAGAATTT GAGACACTGA AAGAGTTTTA CCAGGAAGTG 5340
 CAGAGCTGTA CTGTGACTT AGGTATTACA GCAGACAGCT CCAACCAACC AGACAACAAG 5400
 CACAAGATC GATACATAAA TATCGTTGCC TATGATCATA GCAGGGTTAA GCTAGCACAG 5460
 CTGTCTGAAA AGGATGCAAA ACTGACTGAT TATATCAATG CCAATTATGT TGATGGCTAC 5520
 AACAGACCAA AAGCTTATAT TGCTGCCCAA GGCCCACTGA AATCCACAGC TGAAGATTTG 5580
 TGGAGAATGA TATGGGAACA TAATGTGGAA GTTATTGTCA TGATAACAAA CCTCGTGAG 5640
 AAAGGAAGGA GAAATGTGTA TCAGTACTGG CCTGCGATG GGAGTGAGGA GTACGGGAAC 5700
 TTTCTGTGCA CTCAGAAGAG GTGCAAGTG CTTGCTTATT ATACTGTGAG GAATTTTACT 5760
 CTAAGAAACA CAAAATAAAA AAAGGGCTCC CAGAAAGGAA GACCCAGTGG ACGTGTGGTC 5820
 ACACAGTATC ACTACAGGCA GTGGCCTGAC ATGGGAGTAC CAGAGTACTC CCTGCCAGTG 5880
 CTGACCTTTG TGAGAAAGGC AGCCTATGCC AAGCGCCATG CAGTGGGGCC TGTGTGCTG 5940
 CACTGCACTG CTGGAATGG AAGAACAGGC ACATATATTG TGCTAGACAG TATGTTGCAG 6000
 CAGATTCAAC ACGAAGGAAC TGTCAACATA TTTGGCTTCT TAAACACAT CCGTTACAAA 6060
 AGAAATTAAT TGGTACAAAC TGAGGTGCTG GACAGTCATA TTCTGCTTA TGTTAATGCA 6120
 GCCATACTTA GTAAAGAAAC TGAGGTGCTG GACAGTCATA TTCTGCTTA TGTTAATGCA 6180
 CTCCTCATTC CTGGACAGC AGGCCAAAACA AAGCTAGAGA AACATTTCCA GCTCCTGAGC 6240
 CAGTCAATA TACAGCAGAG TGAATTTCT GCAGCCCTAA AGCAATGCAA CAGGGAAAAG 6300
 AATCGAATT CTCTCATCAT CCCTGTGGAA AGATCAAGGG TTGGCATTTC ATCCCTGAGT 6360
 GGAGAAGGCA CAGACTACAT CAATGCCCTC TATATCATGG GCTATTACCA GAGCAATGAA 6420
 TTCTCATTTA CCCAGCACCT TCTCCTTCAT ACCATCAAGG ATTTCTGGAG GATGATATGG 6480
 GACCAATAG CCCAATGGT GGTATGATT CCTGATGGCC AAAACATGGC AGAAGATGAA 6540
 TTTGTTTACT GGCCAAATAA AGATGAGCCT ATAAATTTGT AGAGCTTTAA GGTCACTCTT 6600
 ATGGCTGAAG AACACAAATG TCTATCTAAT GAGGAAAAAC TTATAATTCA GGACTTTATC 6660
 TTAGAAGCTA CACAGGATGA TTATGTACTT GAAGTGAGGC ACTTTCAGTG TCCTAAATGG 6720
 CCAATCCAG ATAGCCCAT TAGTAAACT TTTGAACCTA TAAGTGTAT AAAAGAAGAA 6780
 GCTGCCAATA GGGATGGGCC TATGATGTT CATGATGAGC ATGGAGGAGT GACGGCAGGA 6840
 ACTTCTGTG CTCTGACAA CTTATGACAC CAACTAGAAA AAGAAAATTC CGTGATGTT 6900
 TACCAGGTAG CCAAGATGAT CAATCTGATG AGGCCAGGAG TCTTGTCTGA CATTGAGCAG 6960
 TATCAGTTTC TCTACAAAGT GATCCTCAGC CTGTGTGACA CAAGGCAGGA AGAGAAATCCA 7020
 TCCACCTCTC TGGACAGTAA TGGTGCAGCA TTGCCTGATG GAAATATAGC TGAGAGCTTA 7080
 GAGTCTTTAG TTTAACACAG AAAGGGGTGG GGGGACTCAC ATCTGAGCAT TGTTTCTCTC 7140
 TTCTTAAAT TAGGCAGGAA AATCAGTCTA GTTCTGTAT CTGTGATTT CCATCACCT 7200
 GACAGTAACT TTCTGACAT AGGATTTCTG OGCCAAATTT ATATCATTA CAATGTGTGC 7260
 CTTTGTGCAA GACTTGTAA TTTACTTATA TGTGAACT AAAATGATTG AATTTTACAG 7320
 TATTTCTAAG AATGGAATTG TGGTATTTT TTCTGTATTT ATTTTACAG AAAATTTCAA 7380
 TTTATAGAGG TTAGGAATTC CAAACTACAG AAAATGTTTG TTTTGTAGT CAAATTTTAA 7440
 GCTGTATTG TAGCAATTAT CAGGTTTGT AGAAATATAA CTTTAAATAC AGTAGCCTGT 7500
 AAATAAACA CTCTTCCATA TGATATTCAA CATTTTACAA CTGCAGTATT CACCTAAAGT 7560
 AGAAATAATC TGTACTTAT TGTAATACT CTGAGTCAAG TTTTCTAGT CTGTGTAAT 7620
 TTTATAATTG TAGATTTTAA TATTTTACTA CTGAGTCAAG TTTTCTAGT CTGTGTAAT 7680
 GTTGTATTG ATGAGTGTG TCAATAGCTG GTCTTACTCT ACCAGTTTTC TGACATTGTA 7740
 TTTGTATTG TAAGTCACTA ACTTGTGTT AGCATGTAAT TTTAACTTT GTGGAAAATA 7800
 GAAATACCTT CATTTTGAAG GAAGTTTATA TGAGATAAC ACCTTACCAA ACATTTGTCA 7860
 AATGGTTTTT ATCCAGGAAA TTGCAAAAAT AAATATAAAT ATTGCCATTA AAAAAAATA 7920
 AAAAAAATA AAAAAAATA AAAA

Seq ID NO: 573 Protein sequence:
Protein Accession #: Eos sequence

5 1 11 21 31 41 51
MRILKRPLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALEQKQNG KKYPCTNSPK 60
QSPINIDEDL TQVNVNLKQL KPQGWDTSL ENTFIRNTGK TVEINLTNDY RVSGGVSEMV 120
10 FKASKITPHW GKCNMSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEAV KGKGLRALS 180
ILFEVGTEN LDFKALIDGV ESVSFRGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240
TDTVDNIVEK DTVSISESQL AVFCEVLTMO QSGVVMIMDY LQNNFREQQY KFSRQVFSY 300
TGKEETHEAV CSSEPEVNOA DPENYTSLLV TWERPRVVYD TMIEKFAVLV QQLDGEDQTK 360
HEFLTDGYQD LGAILNLNLP NMSYVLQIVA ICTNGLYGYK SDQLIVDMPT DNPELDLPFE 420
15 LIGTEEIKE EEEGKDIEEG AIVNPGRDSA TNQIRKKEPQ ISTTHYNRI GTRYNEAKTN 480
RSPTRGSEFS GKGDPVNTSL NSTSQPVTKL ATEKDISLTS QTVELPPHT VEGTSASLND 540
GSKTVLRSPH MNLSGTAESL NTVSITEYEE ESLTSTFKLD TGAEDSSGSS PATSAIFPIS 600
ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDPSEMG NVWFPSTDI 660
TAQPDVGGSR ESFLQNTYTE IRVDESEKTT KSPSAGPVMS QGSPVTDLEM PHYSTFAYFP 720
20 TEVTPHAPTF SSRQDLVST VNVVYSQTQ PVYNGETPLQ PSYSSVFPPL VTPLLLDNQI 780
LNTTPAASVS DSLFHPMPF PVSDVSFESI LSSYDGAPLL PFSSASFSE LFRHLHTVSQ 840
ILPQVTSATE SDKVPLHSL PVAGGDLLE PSLAQYSDVL STTHAASETL EFGSESGVLY 900
KTLMFQVPEP PSSDAMMHAR SSGPEPSYAL SDNEGSQHIF TVSYSSAIPV HDSVGVTYQG 960
SLFSGSPSHIP IPKSSLITPT ASLLQPTHAL SGDGEWSGAS SDSEFLLPDT DGLTALNISS 1020
25 PVSVAEFTYT TSFVGDNDKA LSKSEIYIGN ETELQIPSPN EMVYPSSEST MPNMYDNVVK 1080
LNASLGSTSV SISSTKGMFP GSAHTTTKV FDHEISQVPE NNFVQPTHV VSQASGDTSL 1140
KPVLSANSEP ASSDPASSEM LSPSTQLLFY ETSASFSTEV LLQPSFOASD VDTLLKTLVP 1200
AVSPDPIIVE TPKVDKISST MLHLIVNSNA SSENMLHSTS VPVFDVSPTS HMHSASLQGL 1260
TISYASEKYF PVLLKSESSH QVPSLYSND BLFQTANLEI NQAHPPKGRH VFATPVLSID 1320
30 EPLNTLINKL LHSDEILTST KSSVTGKVFQ GIPTVASDTF VSTDHSPVIG NGHVAITAVS 1380
PHRDGSVSTP KLFLPFSKATP ELSHSAKSDA GLVGGEDGDG TDDGDDDDDD DRGSDGLSIH 1440
KCMSCSSYRE SQEKVMNDSD THENSLMDQN NPISYSLSEN SEEDNRVTSV SSSQTMGDMR 1500
SPGKSPSANG LSQKNDMDKE ENDIQTGSL LPLSPESKAW AVLTSDDEESG SQGQTSDSLH 1560
ENETSTDFS ADTNEKADAG ILAAGDSEIT PGFPQSTSS VTSENSEVH VSEAEASNS 1620
35 HESRIGLAEG LESEKKAIVP LVIVSALTPI CLVLVLGILI YWRKCPQTAH FYLEDSTSPR 1680
VISTPPTPIF PISDDVGAIP IKHFPKHVAD LHASSGFTTE FETLKEFYQE VQSCITDLGI 1740
TADSSNHPDN KHQRYRINIV AYDHSRVKLA QLAEKDGKLT DYINANYVDG YNRPKAYIAA 1800
QGPKLSTAE FWRMIWEHNV EVIVMITNLV EKGRKCDQY WPADGSEBYE NFLVTQKSVQ 1860
VLAYYTVRAF TLRNTKIKKG SQKGRPSGRV VTQYHYTQWP DMGVPEYSLP VLTFRVKAAY 1920
40 AKRHAVGPVW VHCASGVGR GTYIVLDSML QIQIHGHTVN IFGLFKHRS QRNYLVQTEE 1980
QYVFIHDTLV EALLSKETEV LDSHIHAYVN ALLIPGPAGK TKLEKQFOLL SQSNIIQSDY 2040
SAALKQCNRE KNRTSSIIIPV ERSRVGISSL SGEGETDYINA SYIMGYQSN EFIIITQHPLL 2100
HTIKDFWRMI WDHNAQLVVM IPDQGNMAED EFVYVWPKDE PINCESFKVT LMAEHHKCLS 2160
NEEKLIQDQF ILEATQDDYV LEVRHFQCPK WPNPDSPIK TFEIISVIKE EAANRDGPMI 2220
45 VHDEHGGVTA GTFCALTILM HQLEKENSVD VYQVAKMINL MRPGVFPADIE QYQFLYKVL 2280
SLVSTRQEN PSTSLDSNGA ALPDGNIAES LESLV

Seq ID NO: 574 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 148-4518

50 1 11 21 31 41 51
CACACATACG CAGCACGAT CTCACCTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60
CAAAAAAATC ATTCTCTCG CTCCTCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120
55 CGGCGAGGGG CCGCAGACCG TCTGGAATG CGAATCTCAA AGCGTTTCT CGCTTGCAAT 180
CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
CTTGTTGAAG AGATTGGCTG GTCCTATACA GGACCACTGA ATCAAAAAAA TTGGGGAAAG 300
60 AAATATCCAA CATGTAATAG CCAAAACAAA TCTCCTATCA ATATTGATGA AGATCTTACA 360
CAAGTAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAAAC ATCATTGGAA 420
AACACATTCA TGCATAACAC GTGGAAAAACA GTGGAATTA ATCTCACTAA TGACTACCGT 480
GTCAGCGGAG GAGTTTCAGA AATGGTGTTT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540
AAATGCAATA TGTATCTCGA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600
GAGATGCAAA TCTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660
65 GGAAGAGGGA AGTTAAGAGC TTTATCCATT TTGTTGAGG TTGGGACAGA AGAAAATTG 720
GATTTCAAG CGATTATGA TGGAGTCGAA AGTGTAGTC GTTTTGGGAA GCAGGCTGCT 780
TTAGATCCAT TCATCTGTTT GAACCTCTG CCAACTCAA CTGACAAGTA TTACATTTAC 840
AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900
ACAGTTAGCA TCTCTGAAG CGAGTTGGCT GTTTTTGTG AAGTCTTAC AATGCAACAA 960
70 TCTGGTTATG TCATGCTGAT GGACTACTTA CAAAACAATT TTGAGAGCA ACAGTACAAG 1020
TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAGGAAG AGATTATGA AGCAGTTTGT 1080
AGTTACAGAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTTACA 1140
TGGGAAGAGC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGCAGT TTTGTACCAG 1200
CAGTTGGATG GAGAGGACCA AACCAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260
75 GGTGCTATTC TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320
TGCACTAATG TCTCAATATG AAAATACAGC GACCAACTGA TTGTGACAT GCCTACTGAT 1380
AATCCTGAAC TTGATCTTTT CCTGAATTA ATTGGAAGT AAGAAATAT CAAGGAGGAG 1440
GAAGAGGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGCTAGAGA CAGTGCTACA 1500
AACCATAACA GGAAGGAAG AACCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560
80 ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620
AAGGCTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680
ACAGAAAAAG ATATTTCTCT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740
GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAACTG TTCTTAGATC TCCACATATG 1800
AACTTGTGCG GAGTGCAGA ATCCTTAAAT ACAGTTTCTA TAACAGAATA TGAGGAGGAG 1860
AGTTTATGTA CCAATTTCAA GCTTGATATC GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920
85 CCAACTTCTG CTATCCCATT CATCTCTGAG AACATATCCC AAGGATATAT ATTTCTCCTC 1980
GAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AATGCTCTCC 2040
GAAGATTCAA CTTCATCAGG TTCAGAAGAA TCACATAAGG ATCCTTCTAT GGAGGGAAAT 2100

	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACGTAGATA	GGTGTGATG	AATCTGAGAA	GACAAOCCAAG	2220
	TCCTTTCTCG	CAGGCCCACT	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
5	CATTATTCTA	CCTTTGCCTA	CTTCCCACT	GAGSTAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAA	AGGATTGGT	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCCAACCG	2400
	GTATACAAATG	CAGAGGCCAG	TAATAGTAGC	CATGAGTCTC	GTATTGGTCT	AGCTGAGGGG	2460
	TTGGAATCCG	AGAAGAAGGC	AGTTATACCC	CTTGTGATCG	TGTCAGCCCT	GACTTTTATC	2520
	TGCTTAGTGG	TTCTTGTGGG	TATTCTCATC	TACTGGAGGA	AATGCTTCCA	GACTGCACAC	2580
10	TTTACTTAG	AGGACAGTAC	ATCCCTTAGA	GTATATATCA	CACCTCCAAC	ACCTATCTTT	2640
	CCAATTCAG	ATGATGTCGG	AGCAATTTCA	ATAAAGCACT	TTCCAAAGCA	TGTTGCAGAT	2700
	TTACATGCAA	GTAGTGGGTT	TACTGAAGAA	TTTGAGACAC	TGAAAGAGTT	TTACCAGGAA	2760
	GTGCAGAGCT	GTACTGTTGA	CTTAGGTATT	ACACGAGACA	GCTCCAACCA	CCCAGACAAC	2820
	AAGCACAAGA	ATCGATACAT	AAATATCGTT	GCCTATGATC	ATAGCAGGCT	TAAGCTAGCA	2880
	CAGCTTGTCTG	AAAAGGATGC	CAAACTGACT	GATTATATCA	ATGCCAATTA	TGTTGATGGC	2940
15	TACAACAGAC	CAAAAGCTTA	TATTGCTGCC	CAAGGCCAC	TGAAATCCAC	AGCTGAAGAT	3000
	TTCTGGAGAA	TGATATGGGA	ACATAATGTG	GAAGTTATGT	TCATGATAAC	AAACCTCGTG	3060
	GAGAAAGGAA	GGAGAAATG	TGATCAGTAC	TGGCCTGCCG	ATGGGAGTGA	GGAGTACGGG	3120
	AACTTTCTGG	TCACCTCAGAA	GAGTGTGCAA	GTGCTTGCCCT	ATTATCTGT	GAGGAATTTT	3180
	ACTCTAAGAA	ACACAAAAT	AAAAAGGGG	TCCCAGAAAG	GAAGACCCAG	TGGACGTGTG	3240
20	GTCAACAGCT	ATCACTACAC	GCAGTGGCCT	GACATGGGAG	TACCAGAGTA	CTCCCTGCCA	3300
	GTGCTGACCT	TTGCTGGGAA	GGCAGCCTAT	GCCAAGGCC	ATGCAGTGGG	GCCTGTTGTC	3360
	GTCCACTGCA	GTGCTGGAGT	TGGAAGAAC	GGCAGATATA	TTGTGCTAGA	CAGTATGTTG	3420
	CAGCAGATT	AACACGAAGG	AACTGTCAAC	ATATTGGCT	TCTTAAACCA	CATCCGTTCA	3480
	CAAGAAAT	ATTTGGTACA	AACGTAGGAG	CAATATGTCT	TCATTATGTA	TACACTGGTT	3540
25	GAGGCCATAC	TTAGTAAAGA	AACGTAGGAG	CTGGACAGTC	ATATTCTATG	CTATGTTAAT	3600
	GCACCTCTCA	TTGCTGGGAA	AGCAGGCAAA	ACAAAGCTAG	AGAAACAATT	CCAGCTCTG	3660
	AGCCAGTCAA	ATATACAGCA	GAGTGACTAT	TCTGCAGCCC	TAAAGCAATG	CAACAGGGAA	3720
	AAGAATCGAA	CTTCTCTCAT	CATCCCTGTG	GAAAGATCAA	GGGTTGGCAT	TTCTCCCTG	3780
	AGTGGAGAAG	GCACAGACTA	CATCAATGCC	TCCTATATCA	TGGGCTATTA	CCAGAGCAAT	3840
30	GAATTTCATCA	TTACCCAGCA	CCCTCTCCTT	CATACCATCA	AGGATTTCCTG	GAGGATGATA	3900
	TGGGACCATA	TTGCCCAACT	GGTGGTTATG	ATTCCTGATG	GCCAAAACAT	GCCAGAAGAT	3960
	GAATTTGTTT	ACTGGCCCAA	TAAAGATGAG	CCTATAAAT	GTGAGAGCTT	TAAGGTCACT	4020
	CTTATGGCTG	AAGAACACAA	ATGCTATCT	AATGAGGAAA	AACCTATAAT	TCAGGACTTT	4080
	ATCTTAGAAG	CTACACAGGA	TGATTATGTA	CTGAAAGTGA	GGCACTTTCA	GTGCTCTAAA	4140
35	TGGCCAAATC	CAGATAGCCC	CATTAGTAAA	ACTTTTGAAC	TTATAAGTGT	TATAAAGAA	4200
	GAAGCTGCCA	TAGGGATGG	GCCTATGATT	GTTCATGATG	AGCATGGAGG	AGTGACGGCA	4260
	GGAACTTTCT	GTGCTCTGAC	AACCTTATG	CACCAACTAG	AAAAAGAAAA	TTCCGTGGAT	4320
	GTTTACCAGG	TAGCCAAGAT	GATCAATCTG	ATGAGGCCAG	GAGTCTTTGC	TGACATTGAG	4380
	CAGTATCAGT	TTCTCTACAA	AGTGATCCTC	AGCCTTGTGA	GCACAAGGCA	GGAAGAGAT	4440
40	CCATCCACCT	CTCTGGACAG	TAAATGGTGA	GCATTCGCTG	ATGGAAATAT	AGCTGAGAGC	4500
	TTAGAGTCTT	TAGTTTAAAC	CAGAAAGGGG	TGGGGGAGCT	CACATCTGAG	CATTGTTTTC	4560
	CTCTCTCTAA	AATTAGGCAG	GAAATCAGT	CTAGTCTGT	TATCTGTTGA	TTTCCCATCA	4620
	CCTGACAGTA	ACTTTCATGA	CATAGGATTC	TGCCGCCAAA	TTTATATCAT	TAACAATGTG	4680
	TGCCTTTTGT	CAAGACTTGT	AATTTACTTA	TTATGTTTGA	ACTAAATGSA	TGAATTTTA	4740
45	CAGTATTCTT	AAGAATGGAA	TTGTGGTATT	TTTTTCTGTA	TTGATTTTAA	CAGAAAATTT	4800
	CAATTTATAG	AGGTTAGGAA	TTCCAACTA	CAGAAAATGT	TTGTTTGTAG	TGTCAAATTT	4860
	TTAGCTGTAT	TTGTAGCAAT	TATCAGGTTT	GCTAGAAATA	TAACTTTTAA	TACAGTAGCC	4920
	TGTAATAAAA	ACACTCTTCC	ATATGATATT	CAACATTTTA	CAACTGCAGT	ATTCACCTAA	4980
	AGTAGAAATA	ATCTGTTACT	TATTGTAAT	ACTGCCCTAG	TGTCCTCATG	GACCAAAATTT	5040
50	ATATTTATAA	TTGTAGATTT	TTATATTTTA	CTACTGAGTC	AAGTTTCTTA	GTTCGTGTGA	5100
	ATTGTTTATG	TTAATGACCT	AGTTCAATTAG	CTGCTCTTAC	TCTACCAAGT	TTCTGACATT	5160
	GTATTGTGTT	ACCTAAGTCA	TTAATCTTGT	TTAGCATGT	AATTTTAACT	TTTGTGGAAA	5220
	ATAGAAATAC	CTTCAATTTT	AAAGAAGTTT	TTATGAGAA	AACACCTTAC	CAACATTGT	5280
55	TCAAATGGTT	TTTATCCAAG	GAATGCAAAA	AATAAATATA	AATATTGCCA	TTAAAAAAA	5340
	AAAAA	AAAAA	AAAAA				

Seq ID NO: 575 Protein sequence:
Protein Accession #: Eos sequence

60	1	11	21	31	41	51	
	MRILKRFILAC	IQLLCVCRID	WANGVYRQOR	KLVEEIGWSY	TGALNQKNWG	KKYPTCNSPK	60
	QSPINIDEDL	TQVNVNLKKL	KFGWDKTSL	ENTFIHNTGK	TVEINLTNDY	RVSQGVSEMV	120
65	FKASKITFHW	GKCMSSDGS	EHSLEGQKFP	LEMQIYCFDA	DRFSSFEEAV	KGKGLRALS	180
	ILFEVGTEN	LDFKAIIDGV	ESVSRFGKQA	ALDPFILLNL	LPNSTDKYYI	YNGSLTSPPC	240
	TDTVDWIVFK	DTVSISESQL	AVFCEVLTMO	QSGYVLMMDY	LQNNFREQQY	KFSRQVPSSY	300
	TGKEEIHAEV	CSSEPEPVQA	DPENYTSLLV	TWERPRVVD	TMIEKFAVLY	QQLDGEDQTK	360
	HEFLTDGYQD	LGAILNNLLP	NMSYVLQIVA	ICTNGLYKGY	SDQLIVDMPT	DNPELDLPE	420
70	LIGTEELIKE	EBEGKIDIEG	AIVNPGRDSA	TNQIRKKEPQ	ISTTTHYNRI	GTKYNEAKTN	480
	RSPTRGSEFS	KGQDVPNTSL	NSTSQPVTKL	ATEKDISLTS	QTVTELPPHT	VEGTSASLND	540
	GSKTVLRSFH	MNLSGTAESL	NTVSITEVEE	ESLLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIFS	SENPETIYTD	VLIPESARNA	SEDSTSSGSE	ESLKDPMEG	NVWFPSSTDI	660
	TAQPDVGSGR	ESFLQNTYTE	IRVDESEKTT	KSFSAGPVMS	QGPSVTDLDM	PHYSTFAVFP	720
75	TEVTPHAFTP	SSRQDLVST	VNVVVSQTTO	PVYNABEASNS	SHESRIGLAE	GLESEKKAVI	780
	PLVIVSALT	ICLVVLVGLI	IYWRKCFQTA	HFYLEDSTSP	RVISTPPTPI	FPISDVGAI	840
	PIKHFPKHVA	DLHASSGFTE	EPETLKEFYQ	EVQSCVTDLG	ITADSSNHPD	NKHKNRYINI	900
	VAYDHSRVKL	AQLAEKDGKL	TDYINANYVD	GYNRPKAYIA	AQGPLKSTAE	DFWRMIWEHN	960
	VEVIMVITNL	VEKGRKCDQ	YWPADGSEY	GNFLVTQKSV	QVLAAYTVRN	FTLRNTKIKK	1020
80	GSQGRPSGR	VVTQVHYTQM	PDMPVPEYSL	PULTFVRRAA	YAKRHAVGPV	VVHCAGVGR	1080
	TGTYIVLDSM	LQIQHEGTV	NIFGLKHIR	SQRNYLVQTE	EQVFIHDTL	VEAILSKETE	1140
	VLDSEHIHAYV	NALLIPCPAG	KTKLEKQFOL	LSQSNIQQSD	YSAAKQCNR	EKNRTSSSIIP	1200
	VERSRVGISS	LSSEGTDIYN	ASYIMGYQFS	NEFIITQHPL	LHTIKDFWRM	INDHNAQLVV	1260
	NIPDQGNMAE	DEFVYVWPNK	EPINCESFKV	TLMAEEHKCL	SNEEKLIQD	FILEATQDDY	1320
85	VLEVRHFPQP	KWNPDPSPIS	KTFELISVIK	EEAANDRGP	IVHDERGGVT	AGTFPCALITL	1380
	MHOLEKENS	DVYQVAKRIN	LMPGVFADI	EQVQFLYKVI	LSLVSTRQEE	NPSTSLDSNG	1440
	AALPDGNIAE	SLESIV					

Seq ID NO: 576 DNA sequence
Nucleic Acid Accession #: EOS sequence
Coding sequence: 148-4494

5	1	11	21	31	41	51	
	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
10	CGGCGAGGGG	CCGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTGAAG	AGATTGGCTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAA	TGGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCCAAACAA	TCTCCTATCA	ATATGTATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAC	ATCATTTGAA	420
15	AACACATCCA	TTCAATACAC	TGGGAAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTACGCGGAG	GAGTTTCGAG	AATGGTGTTT	AAAGCAAGCA	AGATAACTTT	TCAGTGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCAGAC	CGATTTTCAA	GTTTGTAGGA	AGCAGTCAAA	660
	GGAAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAATTTG	720
20	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTGTGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TGATACTGTT	GAACCTCTCG	CCAACTCAAA	CTGACAAGTA	TTACATTATC	840
	AAATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTTTTGTG	AAGTCTTTAC	ATGCAACAAA	960
	TCTGGTTATG	TCAATGCTGAT	GGACTACTTA	CAAAACCAAT	TTGAGAGCCA	ACAGTACAAG	1020
25	TTCTCTAGAC	AGGTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTTCATGA	AGCAGTTTGT	1080
	AGTTTCAGAC	CAGAAATATG	TCAGGCTGAC	CCAGAGAATT	ATACCAAGCT	TCTTGTACAA	1140
	TGGGAAAGAC	CTCAGTTCGT	TTATGATACC	ATGATTGAGA	AGTTTGCAGT	TTTGTACCAAG	1200
	CAGTTGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATT	TCAATAATTT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
30	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTGACAT	GCCTACTGAT	1380
	AATCCGTAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAATCG	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTTA	AGAAGGGGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAATCA	GGAAAAGGA	ACCCAGATT	TCTACCACAA	CACACTACAA	TGCGATAGGG	1560
35	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTAGTG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACAGTCAC	TAAATTAGCC	1680
	ACAGAAAAG	ATGATCTTCT	GACTTCTCAG	ACTGTGACTG	AACCTGCCACC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAACTCG	TTCTTAGATC	TCCACATATG	1800
	AACTTGTGGG	GGACTGCGAG	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
40	GCAACTCTTG	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCATCAGG	TTCAGAAGAA	TCATAAAGG	ATCCTTCTAT	GGAGGGAAT	2100
	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
45	AGCTTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTGATG	AATCTGAGAA	GACAAACCAAG	2220
	TCTTTTCTG	CAGGCCCACT	GATGTACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CTTTTGGCTA	CTTCCCAACT	GAGGTAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGSATTGGTT	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCAACCG	2400
	GTATACAATG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	2460
50	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
	CTAGTGTTCT	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACAT	COCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	2640
	ATTTTCAGATG	ATGTCGAGGC	AATTCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	2700
	CATGCAAGTA	GTGGGTTTAC	TGAAGAAATT	GAGGAAGTGC	AGAGCTGTAC	TGTTGACTTA	2760
55	GGTATTACAG	CAGACAGCTC	CAACCAACCA	GACAACAAGC	ACAAGAATCG	ATACATAAAT	2820
	ATCGTGGCCT	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	TTGCTGAAAA	GGATGGCAAA	2880
	GCTACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	ACAGACCAAA	AGCTTATATT	2940
	GCTGCCCAAG	GCCCACTGAA	ATCCACAGCT	GAAGATTTC	GGAGAATGAT	ATGGGAACAT	3000
	AATGTGGAG	TTATTGTCAT	GATAACAAC	CTCGTGGAG	AAGGAAGGAG	AAAATGTGAT	3060
60	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAAC	TTCTGGTCAC	TCAGAAGAGT	3120
	GTGCAAGTGC	TTGCCTATTA	TACTGTGAGG	AATTTTACTC	TAAGAAACAC	AAAAATAAAA	3180
	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGA	CGTGGGTCA	CACAGTATCA	CTACAGCGAG	3240
	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	TGACCTTTGT	GAGAAAGGCA	3300
	GCCTATGCCA	AGGCCCATGC	AGTGGGGCCT	GTTGTCGTCC	ACTGCAAGTG	TGGAGTTGGA	3360
	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	AGATTCAACA	CGAAGGAAC	3420
65	GTCAACATAT	TTGGCTTCTT	AAAAACATC	CGTTCAACAA	GAATATTATT	GGTACAAACT	3480
	GAGGAGCAAT	ATGTCTTCAT	TCATGATACA	CTGGTTGAGG	CCATCTTAGG	TAAAGAAACT	3540
	GAGGTGCTGG	ACAGTCATAT	TCATGCCTAT	GTTAATGCAC	TCCTCATTCC	TGGACGAGCA	3600
	GGCAAAACAA	AGCTAGAGAA	ACAATTCAG	CTCCTGAGCC	AGTCAAATAT	ACAGCAGAGT	3660
	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAAAGA	ATCGAACTTC	TTCTATCATC	3720
70	CCTGTGGAAA	GATCAAGGGT	TGGCATTTC	TCCTGAGTG	GAGAAGGCAC	AGACTACATC	3780
	AATGGCTCCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	TCATCATTAC	CCAGCACCTC	3840
	CTCCTTCATA	CCATCAAGGA	TTCTGGAGG	ATGATATGGG	ACCATAATGC	CCAACTGGTG	3900
	GTTATGATTC	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	TTGTTTACTG	GCCAAATAAA	3960
75	GATGAGCCTA	TAAATTGTGA	GAGCTTTAAG	GTCACCTTTA	TGGCTGAAGA	ACACAAATGT	4020
	CTATCTAATG	AGGAAAAACT	TATAATTGAG	GACTTTTATC	TAGAAGCTAC	ACAGGATGAT	4080
	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	CAAACTCAGA	TAGCCCATTT	4140
	AGTAAACTTT	TTGAACCTAT	AAGTGTATTA	AAAGAAAGAG	CTGCCAATAG	GGATGGGCTC	4200
	ATGATTGTTT	ATGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	CTTTCTGTGC	TCTGACAACC	4260
80	CTTATGCAAC	AACAGAAAAA	AGAAAATTCC	GTGGATGTTT	ACCAGGTAGC	CAAGATGATC	4320
	AATCTGATGA	GGCCAGGAGT	CTTTGCTGAC	ATTGAGCAGT	ATCAGTTTCT	CTACAAAGTG	4380
	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	CCACCTCTCT	GGACAGTAAT	4440
	GGTGAGCAT	TGCTGATGG	AAATATAGCT	GAGAGCTTAG	AGTCTTTAGT	TTAACACAGA	4500
	AAGGGGTGGG	GGGACTCACA	TCTGAGCATT	GTTTTCTCT	TCCTAAATTT	AGGCAGGAAA	4560
	ATCAGTCTAG	TTCTGTATC	TGTTGATTTC	CCATCACCTG	ACAGTAACCT	TCATGACATA	4620
85	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	TTTTTGCAAG	ACTTGTAAAT	4680
	TACTTATTAT	GTTTGAACTA	AAATGATTGA	ATTTTACAGT	ATTTCTAAGA	ATGGAATTGT	4740
	GGTATTTTTT	TCGTATTGTA	TTTTAACAGA	AAATTTCAAT	TTATAGAGGT	TAGGAATTCC	4800

WO 02/086443

AAACTACAGA AAAATGTTTGT TTTTAGTGT CAAATTTTGT AGCAATTATC 4860
 AGGTTTGCTA GAAATATAAC TTTTAATACA GTAGCCTGTA AATAAACAC TCTTCCATAT 4920
 GATATTCAAC ATTTTACAAC TGCAGTATTC ACCTAAAGTA GAAATAATCT GTTACTTATT 4980
 GTAAATACG CCCTAGTGTG TCCATGGACC AAAATTATAT TTATAATTGT AGATTTTAT 5040
 ATTTTACTAC TGAGTCAAGT TTTCTAGTTC TGTGTAATTG TTTAGTTTAA TGACGTAGTT 5100
 CATTAGCTCG TCTTACTCTA CCAGTTTCT GACATTGTAT TGTGTTACCT AAGTCATTAA 5160
 CTTTGTTC CAATGTAAIT TTAACTTTG TCGAAATAG AAATACCTTC ATTTTGAAAG 5220
 AAGTTTATAT GAGAATAACA CCTTACCAA CATGTGTTCA ATGGTTTATA TCCAAGGAAT 5280
 TCGAAAAATA AATATAATA TTGCCATTAA AAAAAAAAAA AAAAAAAAAA 5340
 AAA

Seq ID NO: 577 Protein sequence:
 Protein Accession #: EOS sequence

1 11 21 31 41 51
 MRILKRFLAC IQLLCVCRLD WANGYYRQOR KLVEEIGWSY TGAUNQKNWG KKYPTCNSPK 60
 QSPINIDEDL TQVNVNKKL KFQGWDKTSL ENTPIHNTGK TVEINLNDY RVSGGVSEMV 120
 FKASKITFW GKCMSSDGS EHSLEGQKFP DEMQYCFDA DRFSSFEEAV KKGKGLRLALS 180
 ILFEVGTEN LDFKAIIDGV ESVSFRGKQA ALDPFILLNL LFNSTDKYYI YNGSLTSPPC 240
 TDTVDWIVFK TDTVSIESQL AVFCEVLTMO QSGYVLMMDY LQNNFREQQY KFSRQVFSY 300
 TGKEEIEHAV CSSEPENVOA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360
 HEFLTIDGYQD LGAILNLLP NMSYVLQIVA ICTNGLYGYK SDQLIVDMPT DNPDLDFPE 420
 LGITEEIIEE EEEGKDIIEG AIVNPGRDSA TNQIRKKEPQ ISTTHYNRI GTKYNEAKTV 480
 RSPTRGSEFS GKGDPNTSL NSTSQPVTKL ATEKDISLTS QTVTELPPTT VEGTSASLND 540
 GSKTVLRSPH NLHSSGTESL NTVSITTEYE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600
 ENISQGYIFS SENPETITVD VLIPESARNA SEDSTSSGSE ESLKDPSEMG NVWFPSSTDI 660
 TAQPDVGSGR ESFLQNTYTE IRVDESEKTT KSFSAGPVMS QGPSVTDEM PHYSTFAYFP 720
 TEVTPHAFPT SSRQDLVST VNVVYSQTTQ PVYNEASNS HESRIGLAEG LESEKKAVIP 780
 LVIVSALTPI CLVVLVGLIL YWRKCFQTAH FYLEDSTSPR VISTPPTPIE PISDDVGAIP 840
 IKHPPKHVAD LHASGGTTEE FEEVQCTVD LGITADSNH PDNKHKNRYI NIVAYDHSRV 900
 KLAQLAEKDG KLTDIYINANY VDGYNRPKAY IAAQGPKLST AEDFWRMIWE HNVFVIMIT 960
 NLVEKGRKRC DQYWPADGSE EYGNFLVTQK SVQVLAYYTV RNFTLRNTKI KKGSKGRPS 1020
 GRVVTQYHYT QWPMGVPEY SLPLVTFVRK AAYAKRHAVG PVVHCSAGV GRTGTIYVLD 1080
 SMLQIQHESG TVNIFGLKH IRSQRNYLVQ TEEQYVFIHD TLVEAILSKE TEVLDSHHA 1140
 YVNALLPFG AKTKLEKQF QLSQSNIOQ SDYSAALKQC NREKNRTSSI IPVRSRVGI 1200
 SSSLGEGTDY INASYIMGY QSNFIIQTH PLLHTIKDFW RMIDHNAQL VVMIPDQNM 1260
 AEDEFYVWPN KDEPINCESF KVTLMABEHK CLSNEEKLI QDFILEATQD DYLVEVRHFQ 1320
 CPKMPNDSP ISKTFELISV IKEEAANRDG PMIVDEHGG VTAGTFPCALT TLMHLEKEN 1380
 SVDVYQVAKM INLMRPGVFA DIEQYQFLYK VILSLVSTRO EENPSTSLDS NGAALPDGNI 1440
 AESLESIV

Seq ID NO: 578 DNA sequence
 Nucleic Acid Accession #: EOS sequence
 Coding sequence: 501-4514

1 11 21 31 41 51
 CACACATACG CACGACGAT CTCACCTGGA TCTATACACT GGAGGATTAA AACAAACAAA 60
 CAAAAAACA ATTTCCTCTG CTCCTCCCTC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120
 CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCTCAA AGCGTTTCCT CGCTTGCAAT 180
 CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
 CTTGTTGAAG AGATTGGCTG GTCCATACAC GGAGCACTGA ATCAAAAAAT TGGGAAAGA 300
 AATATCCAAC GAATTAATAG CCAAAACAAT CTCCTATCAA TATTGATGAA GATCTTACAC 360
 AAGTAAATFT GATCTTAAG AAACCTAAAT TTCAGGGTTG GGATAAACA TCATTGGAAA 420
 ACACATTCAAT TCATAACACT GGGAAACAG TGGAAATTA TCTCACTAAT GACTACCGTG 480
 TCAGCGGAGG AGTTTCAGAA ATGCTGTTTA AAGCAAGCAA GATAACTTTT CACTGGGGA 540
 AATGCAATAT GTCATCTGAT GGATCAGAGC ATAGTTTAGA AGGACAAAAA TTTCCACTTG 600
 AGATGCAAT CTACTGCTTT GATGCGGACC GATTTCAGG TTTTGAGGAA GCAGTCAAG 660
 GAAAAGGGAA GTTAAGAGCT TTATCCATTT TGTTTGAGGT TGGACAGAAA GAAAATTTGG 720
 ATTTCAAGGC GATTATTGAT GGAGTCGAAA GTGTTAGTCG TTTTGGGAGG CAGGCTGCTT 780
 TAGATCCATT CATACTGTTG AACCTTCTGC CAACTCAAC TGACAAGTAT TACATTTACA 840
 ATGGCTCATT GACATCTCCT CCTGACAG ACACAGTTGA CTGGATTGTT TTTAAAGATA 900
 CAGTTAGCAT CTCTGAAAGC CAGTTGGCTG TTTTGTGTA AGTTCTTACA ATGCAACAAT 960
 CTGTTATGTT CATGCTGATG GACTACTTAC AAAACAATTT TCGAGAGCAA CAGTACAGT 1020
 TCTCTAGACA GGTGTTTTC TCATACACTG GAAAGGAAGA GATTATGAA GCAGTTTGTA 1080
 GTTCAGAAC AGAAAAATGTT CAGGCTGACC CAGAGAATTA TACCAGCCTT CTGTTACAT 1140
 GGGAAAGACC TCGAGTCGTT TATGATACCA TGATTGAGAA GTTTGCAGTT TGTACCAGC 1200
 AGTTGGATGG AGAGGACCAA ACCAAGCATG AATTTTGTAC AGATGGCTAT CAAGACTTGG 1260
 GTGCTAATCT CAATAATTGT CTACCCAATA TGAGTTATGT TCTTCAGATA GTAGCCATAT 1320
 GCACTAATGG CTTATATGGA AATACAGCG ACCAAGTATG TGTGACATG CTAATGATA 1380
 ATCTGAACT TGATCTTTTC CCTGAATTA TTGGAATTA AGAATAATC AAGGAGGAGG 1440
 AAGAGGGAAA AGACATTGAA GAAGGCGCTA TTGTGAATCC TGGTAGAGAC AGTGCTACAA 1500
 ACCAATCAG GAAAAGGAAA CCCCAGATTT CTACCAACAC ACCTACAAT CGCATAGGGA 1560
 CGAATACAA TGAAGCCAAG ACTAACCGAT CCCCACAAG AGGAAGTGAA TTCTCTGAAA 1620
 AGGGTGATGT TCCCAATACA TCTTTAAATT CCACTCCCCA ACCAGTCACT AAATAGCCA 1680
 CAGAAAAGA TATTTCCCTG ACTTCTCAGA CTGTGACTGA ACTGCCACCT CACACTGTGG 1740
 AAGGTACTTC AGCCTCTTTA AATGATGGCT CTAAGACTGT TCTTAGATCT CCACATATGA 1800
 ACTTGTGGGG GACTGCGAGAA TCTTAAATA CAGTTTCTAT AACAGAATAT GAGGAGGAGA 1860
 GTTTATTGAC CATTGTTCAAG CTGATACTG GAGCTGAAGA TTCTTCAGCG TCCAGTCCCG 1920
 CACTTCTGCT TATCCCATTC ATCTCTGAGA ACATATCCCA AGGATATATA TTTTCTCCG 1980
 AAAACCCAGA GACATAACA TATGATGTC TATACCAGA ATCTGCTAGA AATGCTTCCG 2040
 AAGATTCAAC TTCATCAGGT TCAGAAAGAA CACTAAAGGA TCTTCTATG GAGGGAATG 2100
 TGTGGTTTCC TAGCTCTACA GACATAACAG CACAGCCCGA TGTGGATCA GGCAGAGAGA 2160
 CCTTCTTCCA GACTAATTAC ACTGAGATAC GTGTTGATGA ATCTGAGAAG ACAACCAAGT 2220
 GCTTTCTGCT AGGCCAGTG ATGTACAGG GTCCCTCAGT TACAGATCTG GAAATGCCAC 2280

5	ATTATTCTAC	CTTTGGCTAC	TTCCCACTG	AGGTAACACC	TCATGCTTTT	ACCCCATCCT	2340
	CCAGACAACA	GGATTGGTG	TCCACGGTCA	ACGTGGTATA	CTCGCAGACA	ACCCAACCGG	2400
	TATACAATGA	GGCCAGTAAT	AGTAGCCATG	AGTCTCGTAT	TGGTCTAGCT	GAGGGGTGG	2460
	AATCCGAGAA	GAAGGCGATT	ATACCCCTTG	TGATCGTGTC	AGCCCTGACT	TTTATCTGTC	2520
	TAGTGGTTCT	TGTGGGTATT	CTCATCTACT	GGAGGAAATG	CTTCCAGACT	GCACACTTTT	2580
	ACTTAGAGGA	CAGTACATCC	CCTAGAGTTA	TATCCACACC	TCCAACACCT	ATCTTTCCAA	2640
	TTTCAGATGA	TGTGGGAGCA	ATTCCAATAA	AGCACTTTCC	AAAGCATGTT	GCAGATTTAC	2700
	ATGCAAGTAG	TGGGTTTACT	GAAGAAATTTG	AGACACTGAA	AGAGTTTAC	CAGGAAGTGC	2760
10	AGAGCTGTAC	TGTTGACTTA	GGTATTACAG	CAGACAGCTC	CAACCAACCA	GACAAACAGC	2820
	ACAAGAAATCG	ATACATAAAT	ATCGTTGCCT	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	2880
	TTGCTGAAAA	GGATGGCAAA	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	2940
	ACAGACCAAA	AGCTTATATT	GCTGCCCAAG	GCCCACTGAA	ATCCACAGCT	GAAGATTTCT	3000
	GGAGAATGAT	ATGGGAACAT	AATGTGGAAG	TTATTGTCAT	GATAACAAAC	CTCGTGAGAG	3060
	AAGGAAGGAG	AAAAATGTGAT	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAAGT	3120
15	TTCTGGTCA	TCAGAGAGAGT	GTGCAAGTGC	TTGCCTATTA	TACTGTGAGG	AATTTTACTC	3180
	TAAGAAACAC	AAAAATAAAA	AAGGCTCTCC	AGAAAGGAAG	ACCCAGTGGG	CGTGTGGTCA	3240
	CACAGTATCA	CTACACGAG	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	3300
	TGACCTTTGT	GAGAAAGGCA	GCCTATGCCA	AGCGCCATGC	AGTGGGCGCT	GTTCGCTGCC	3360
	ACTGCACTGC	TGGAGTTGGA	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	3420
20	AGATTCAACA	CGAAGGAAGT	GTCAACATAT	TTGGCTTCTT	AAAAACATC	CGTTTCAAAA	3480
	GAATTTATT	GGTACAAACT	GAGGAGCAAT	ATGCTTTCAT	TCATGATACA	CTGGTTGAGG	3540
	CCATCTACTG	TAAAGAAACT	GAGGTGCTGG	ACAGTCAAT	TCATGCCCTAT	GTAAATGCAC	3600
	TCCTCATTC	TGGACCAAGC	GGCAAAACAA	AGCTAGAGAA	ACAATTCAG	CTCCTGAGCC	3660
	AGTCAAAATAT	ACAGCAGAGT	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAGAAG	3720
25	ATCGAACTTC	TTCTATCATC	CCTGTGGAAA	GATCAAGGGT	TGGCATTTCA	TCCCTGAGTG	3780
	GAGAAGGCAC	AGACTACATC	AATGCTCTCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	3840
	TCATCATTC	CCAGCACCCT	CTCCTTCATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	3900
	ACCATATGTC	CCAATCTGGT	GTTATGATTC	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	3960
30	TTGTTTACTG	GCCAAATAAA	GATGAGCCTA	TAAATTGTGA	GAGCTTAAAG	GTCACTCTTA	4020
	TGGCTGAAGA	ACACAAATGT	CTATCTAATG	AGGAAAAACT	TATAATTCAG	GACTTTATCT	4080
	TAGAAGCTAC	ACAGGATGAT	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	4140
	CAATTCAGA	TAGTCCCTACT	AGTAAACTT	TTGAACCTAT	AAGTGTATA	AAAGAAGAAG	4200
	CTGCCAATAG	GAGTGGGCTT	ATGATTGTTT	ATGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	4260
35	TTTCTGTGTC	TCTGACAACC	CTTATGCACC	AACAGAAAA	AGAAAAATTC	GTGGATGTTT	4320
	ACCAGGTAGC	CAAGATGATC	AATCTGATGA	GGCCAGGAGT	CTTTCCTGAC	ATTGAGCAGT	4380
	ATCAGTTTCT	CTACAAAGTG	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	4440
	CCACCTCTCT	GGACAGTAAT	GGTGAGCAT	TGCTGATGG	AAATATAGCT	GAGAGCTTAG	4500
	AGTCTTTAGT	TTAACAAGA	AAGGGGTGGG	GGGACTCACA	TCTGAGCATT	GTTCCTCTCT	4560
40	TCCTAAATAT	AGGCAGGAAA	ATCAGTCTAG	TTCTGTTATC	TGTTGATTTC	CCATCACCTG	4620
	ACAGTAACCT	TCATGACATA	GGATCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	4680
	TTTTTGCAAG	ACTTGTAAT	TACTTATAT	GTTTGAACCTA	AAATGATTGA	ATTTTACAGT	4740
	ATTCTAAGA	ATGGAATTTG	GGTATTTT	TCTGTATTGA	TTTTAACAGA	AAATTTCAAT	4800
	TTATAGAGGT	TAGGAATTTCC	AAACTACAGA	AAATGTTTGT	TTTTAGTGT	AAATTTTAG	4860
45	CTGTATTTGT	AGCAATTATC	AGGTTTGCTA	GAATATAAC	TTTTAATACA	GTAGCCTGTA	4920
	AATAAAACAC	TCTTCCATAT	GATATTCAAC	ATTTTACAAC	TGCAGTATTC	ACCTAAAGTA	4980
	GAATAATCT	GTTACTTATT	GTAATACTG	CCCTAGTGTC	TCCATGGACC	AAATTTATAT	5040
	TATATAATGT	AGATTTTAT	ATTTTACTAC	TGAGTCAAGT	TTTCTAGTTC	TGTGTAATG	5100
	TTTAGTTTAA	TGAGCTAGTT	CATTAGCTGG	TCTTACTCTA	CCAGTTTCTT	GACATTGTAT	5160
50	TGTGTTACCT	AAGTCATTAA	CTTGTGTTCA	GCATGTAATT	TTAATCTTTG	TGGAATAAG	5220
	AAATACCTTC	ATTTTGAAG	AAGTTTTTAT	GAGAATAACA	CCTTACCAAA	CATTGTTCAA	5280
	ATGGTTTTTA	TCCAAGGAAT	TGCAAAATA	AATATAAATA	TTGCCATTAA	AAAAAAAAAA	5340
	AAAAAAAAAA	AAAAAAAAAA	AAA				

Seq ID NO: 579 Protein sequence:
Protein Accession #: EOS sequence

55	1	11	21	31	41	51	
	MVFKASKITP	HWGKCNMSSD	GSEHSLEGQK	FPLEMQIYCF	DADRFSSPFE	AVKGGKGLRA	60
60	LSILFEVGTG	ENLDFKAIID	GVESVSRPGK	QAALDPFILL	NLLPNSTDKY	YIYNGSLTSP	120
	PCTDVTVDIV	PKDTVSISES	QLAVFCEVLT	MQQSGYVLM	DYLNQNFREQ	QYKFSRQVFS	180
	SYTGKEIHE	AVCSSEPENV	QADPENYTSI	LVTWERPRVV	YDTMIEKFAV	LYQLDGEDQ	240
	TKHEFLTDGY	QDLGAILNNL	LPNMSYVLQI	VAICTNGLYG	KYSDQLIVDM	PTDNPELDLF	300
65	PELIGTEBII	KEEEEGKIDIE	EGAIVNPGRD	SATNQIRKKE	PQISTTHYIN	RIGTKYNEAK	360
	TNRSPTRGSE	FSGGKDVPEVT	SLNSTSQPVT	KLATEKDISL	TSQVTELPFP	HTVEGTSASL	420
	NDGSKTVLRS	PHMNLSTGAE	SLNTVSITEY	EEESLLTSFK	LDTGAEDSSG	SSPATSAIPP	480
	ISENISQGYI	FSSNPETIT	YDVLIPESAR	NASEDSTSSG	SEESLKDPSP	EGNVWFPST	540
	DITAPQDVGS	GRESFLQTNV	TEIRVDESEK	TTKFSAGPV	MSQGPSVTDL	EMPHYSTFAY	600
70	FPTEVTPHAF	TPSSRQQLDV	STVNVVYSQT	TQPVYNEASN	SSHESRIGLA	EGLESEKKAV	660
	IPLVIVSALT	FICLVVLVGI	LIYWRKCFQT	AHFYLEDSTS	PRVISTPPTP	IFPISDDVGA	720
	IPIKHPKPVK	ADLHASSGFT	EEFETLKEFY	QEVQSCITVDL	GITADSSNHP	DNKHQRYIN	780
	IVAYDHSRVK	LAQLAEKDGK	LTDYINANYV	DGYNRPKAYI	AAQGPLKSTA	EDFWRMINEH	840
	NVEVIVMITN	LVEKGRKCD	QYWPADGSEE	YGNFLVTQKS	VQVLAAYTVR	NFTLRNTKIK	900
75	KGSQKGRPSG	RVVTQYHYTO	WPDMGVPEYS	LPVLTFRVKA	AYAKRHVGP	VVVHCSAGVG	960
	RTGTYIVLDS	MLQIQIHGEGT	VNIFGFLKHI	RSQRNYLVQT	EEQYVFIHDT	LVEAILSKET	1020
	EVLDSHIIHAY	VNALLIPGPA	GKTKLEKQFQ	LSSQSNIQQS	DYSAALKQCN	REKNTSSII	1080
	PVERSRVGIS	SLSGEGTDYI	NASYIMGYO	SNEFIITQHP	LLHTIKDFWR	MIWDHNAQLV	1140
	VMIPGQDMA	DEFFVYWNK	DEPINCESFK	VTLMABEHKC	LSNEEKLIQ	DFILEATQDD	1200
80	YVLEVRHFQC	PKWPNPDSPI	SKTFELISVI	KEEAANRDGP	MIVHDEHGGV	TAGTFCALTT	1260
	LMHQLKENS	VDVYQVAKMI	NLMRPGVFAD	IEQYQFLYKV	ILSLVSTRQE	ENPSTSLDSN	1320
	GAALPDGNIA	ESLESIV					

Seq ID NO: 580 DNA sequence
Nucleic Acid Accession #: EOS sequence
Coding sequence: 148-4632

1	11	21	31	41	51
---	----	----	----	----	----

	CACACATAAG	CACGCAGCAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCCTTGG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGCGAGGGGG	CCGCAGACCG	TCTGGAAATG	CGAATCTTAA	AACGTTTCCT	CGCTTGCAAT	180
5	CAGCTCCTCT	GTGTTTGGCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCCATATACA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAA	300
	AAATATCCAA	CATGTAAATAG	CCCCAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420
	AACACATTCA	TTCAATACAC	TGGGAAACAA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
10	GTCAGCGGAG	GAGTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCAGTGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGSATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTGTAGGA	AGCAGTCAAA	660
	GGAAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAAATTG	720
	GATTTCAAAG	CGATTATTTA	TGGAGTCGAA	AGTGTTAGTC	GTTTGGGGAA	GCAGGCTGCT	780
15	TTAGATCCAT	TCATACTGTT	GAACCTTCTG	CCAACTCAA	CTGACAAGTA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAGG	CCAGTTGGCT	GTTTTTTTGT	AAGTCTTCTA	AATGCAACAA	960
	TCTGGTTATG	TCAATCTGAT	GGACTACTTA	CAAAACAATT	TTGAGAGACA	ACAGTACAG	1020
20	TTCTCTAGAC	AGGTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTCTATGA	AGCAGTTTGT	1080
	AGTTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTATAC	1140
	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGCAGT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGTG	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATT	TCAATAATT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
25	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAACCT	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGTACAA	1500
	AACCAATACA	GGAAAAAGGA	ACCCACAGAT	TCTACCACAA	CACACTACAA	TGCTATAGG	1560
	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
30	AAGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCAAGTAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTCTCTT	GACTTCTCAG	ACTGTGACTG	AAGTGCACCC	TCACACTGTG	1740
	GAAGGTACTT	CAGGCTCTTT	AAATGATGGC	TCTAAAACCT	TTCTTAGATC	TCCCATATAT	1800
	AACCTTGTGG	GGAGCTGCAG	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTTA	CGAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
35	GCAACTTCTG	CTATCCCATC	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAAACCCAG	AGACAATAAT	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	TTCTCATCAG	TTCAAGAGAA	TCACAAAGG	ATCCTTCTAT	GGAGGGAAT	2100
	GTGTGGTTTT	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
40	AGCTTCTCTC	AGACTAATTA	CACTGAGATA	CGTGTGTATG	AATCTGAGAA	GACACCCAG	2220
	TCTTTTCTG	CAGGCCAGT	GATGTACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CCTTTGCCTA	CTTCCCAACT	GAGGTAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTTTGG	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCCAACC	2400
	GTATACAAAT	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGTTTG	2460
45	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAAACCC	TATCTTTCCA	2640
	ATTTAGATG	ATGTCGGAGC	AATTTCAATA	AAGCACITTC	CAAGCATGT	TGCAGATTTA	2700
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTTA	CCAGGAAGTG	2760
50	CAGAGCTGTA	CTGTTGACTT	AGGTATTACA	CGAGACAGCT	CCAACCAACC	AGACAACAG	2820
	CACAAGAAAT	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	2880
	CTTGCTGAAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	2940
	AACAGACCAA	AGGCTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTC	3000
	TGGAGAAATG	TATGGGAACA	TAATGTGGAA	GTTATTGTCA	TGATAACAAA	CCTCGTGAG	3060
	AAAGGAAGGA	GAAATGTGTA	TCAGTACTGG	CCTGCGGATG	GGAGTGAGGA	GTACGGGAAC	3120
55	TTTCTGTGTA	CTCAAGAGAG	TGTGCAAGTG	CTTGCCATT	ATACTGTGAG	GAATTTTACT	3180
	CTAAGAAACA	CAAAATTAAG	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	3240
	ACACAGTATC	ACTACAGCCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	3300
	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGCCCATG	CAGTGGGGCC	TGTTGTCTCT	3360
	CAGTCAAGT	CTGGAGTTGG	AAGAACAGGC	ACATATATTG	TGCTAGACAG	TATGTTGCAG	3420
60	CAGATTCAAC	ACGAAGGAAC	TGTCACACATA	TTTGGCTTCT	TAAACACAT	CCGTTACAAA	3480
	AGAAATTTAT	TGTAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTCATGCCTA	TGTTAATGCA	3540
	GCCATACTTA	GTAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTCATGCCTA	TGTTAATGCA	3600
	CTCCTCATTC	CTGGACCAAG	AGGCAAAACA	AAGCTAGAGA	AACAATTCCA	GGGTCTCACT	3660
	CTGTCAACCA	GGCTGGAGTG	CAGAGGCACA	ATCTGGCTC	ACTGCAACCT	TCTCTCCCT	3720
65	GGCTTAACTG	ATCCTCTCTC	CTCAGCCTCC	CGAGTGGCTG	GGACTATACT	CCTGAGCCAG	3780
	TCAATATATC	ACGAGAGTGA	CTAATTCTGCA	GGCCTAAAGC	AATGCAACAG	GGAAAGAAAT	3840
	CGAACTTCTT	CTATCATCCC	TGTGGAAAGA	TCAAGGGTTG	GCATTTCATC	CCTGAGTGGA	3900
	GAAGGCACAG	ACTACATCAA	TGCCCTCTAT	ATCATGGGCT	ATTACCAGAG	CAATGAATTC	3960
	ATCATTAACC	AGCACCTCTC	CCTTCATACC	ATCAAGGATT	TCTGGAGGAT	GATATGGGAC	4020
70	CATAATGCC	AACTGGTGGT	TATGATTCCT	GATGGCCAAA	ACATGGCAGA	AGATGAATTT	4080
	GTTTACTGGC	CAAAATAAGA	TGAGCCTATA	AATTGTGAGA	GCTTAAAGGT	CACCTTATG	4140
	GCTGAAGAAC	ACAAATGTCT	ATCTAATGAG	GAAAACTTAA	TAATTGAGGA	CTTTATCTTA	4200
	GAAGCTACAC	AGGATGATTA	TGTACTTGAA	GTGAGGCACT	TTCAAGTCTC	TAAATGGCCA	4260
	AATCCAGATA	GCCCCATTAG	TAAAACTTTT	GAACTTATAA	GTGTTATAAA	AGAAGAAGCT	4320
75	GGCAATAGGG	ATGGGCTGAT	GATTGTTTCT	GATGAGCATG	GAGGAGTGAC	GGCAGGAAC	4380
	TTCTGTGCTC	TGACAACTCT	TATGACCAAA	CTAGAAAAAG	AAAATTCCTG	GGATGTTTAC	4440
	CAGGTAGCCA	AGATGATCAA	TCTGATGAGG	CCAGGAGTCT	TTGCTGACAT	TGAGCAGTAT	4500
	CAGTTTCTCT	ACAAAGTGAT	CCTCAGCCTT	GTGGGACAAA	GGCAGGAAGA	GAATCCATCC	4560
	ACCTCTCTGG	ACAGTAATGG	TGCAGCATTG	CCTGATGGAA	ATATAGCTGA	GAGCTTAGAG	4620
80	TCTTTAGTTT	AACACAGAAA	GGGGTGGGGG	GACTCACATC	TGAGCATTGT	TTTCTCTTTC	4680
	CTAAAATTAG	GCAGGAAAAA	CAGTCTAGTT	CTGTTATCTG	TTGATTCTCC	ATCACTGAC	4740
	AGTAACITTC	ATGACATAGG	ATTCTGCCGC	CAAAATTATA	TCATTAAACA	TGTGTGCTT	4800
	TTTGCAAGAC	TTGTAATTTA	CTTATTATGT	TTGAACATAA	ATGATTGAAT	TTTACAGTAT	4860
	TTCTAAGAA	GGAAATTGGG	TATTTTCTTC	TGTATTGATT	TTAACAGAAA	ATTTCAATTT	4920
	ATAGAGGTTA	GGAAATCCAA	ACTACAGAAA	ATGTTTGTTC	TTAGTGTCAA	ATTTTGTAGCT	4980
85	GTATTTGTAG	CAATTATCAG	GTGCTGCTAG	AATATAACTT	TTAATACAGT	AGCCTGTAAA	5040
	TAAACACTC	TTCCATATGA	TATTAACACT	TTTACACTG	CAGTATTCAC	CTAAGTAGA	5100
	AATAATCTGT	TACTTATTGT	AAATACTGCC	CTAGTGTCTC	CATGGACCAA	ATTTATATTT	5160

ATAATTGTAG ATTTTATAT TTTACTACTG AGTCAAGTTT TCTAGTCTG TGTAAATTGT 5220
 TAGTTTAATG ACGTAGTTCA TTAGCTGGTC TTACTCTACC AGTTTTCTGA CATTGTATTG 5280
 TGTACTCTAA GTCAATTAAT TTGTTTCAGC ATGTAATTTT AACTTTTGTG GAAAAATAGA 5340
 ATACCTTCAT TTTGAAAGAA GTTTTATGA GAATAACACC TTACCAACA TGTTCAAAT 5400
 5 GGTTTTATC CAAGGAATTG CAAAAATAA TATAATATT GCCATTAATA AAAAAAATAA 5460
 AAAAAAAAAA AAAAAAAAAA A

Seq ID NO: 581 Protein sequence:
 Protein Accession #: EOS sequence

1 11 21 31 41 51
 MRILKRLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALEQKNWG KKYPTCNPK 60
 QSPINIDEDL TQVNVNLKGL KFGWDKTSI ENTFIHNTGK TVEINLNDY RVSGGVSEMV 120
 15 FKASKITFW GKCMNSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEAV KKGKLRALS 180
 ILFEVGTEN LDFKAIIDGV ESVSFRFGQA ALDPPILLNL LPNSTDKYI YNGSLTSPPC 240
 TDTVDWIVFK DTVSISESQL AVFCEVLTMO QSGYVMLMDY LQNNFREQQY KFSRQVSSY 300
 TGKEEIHFAV CSSEPNVQA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360
 HEFLTQGYQD LGAILNNLLP NMSYVLQIVA ICTNGLYKGY SDQLIVDMPT DNPDLDFPE 420
 20 LIGTEEIIKE EEEKDIIEG AIVNPRDSA TNQIRKKEPO ISTTHYNRI GTKYNKAKTN 480
 RSPTRGSEFS GKGDVNTSL NSTSQPVTKL ATEKDILSTS QTVTELPHPT VEGTSASLND 540
 GSKTVLRSPH NMLSGTAESL NTVSITYEYB ESLTSFKLD TGAEDSSGSS PATSAIPFIS 600
 ENISQGIYS SENPETITYE VLIPESARNA SEDSTSSGSE ESLKDPSEMG NWVFPSTDI 660
 25 TAQPDVGSGR BSFLQNTYD IRVDESEKTT KSFSAGPVMS QGSPVTDLEM PHYSTFAFFP 720
 TEVTPHAFPT SSRQDLVST VNVVYSQTTO PVYNEASNS HESRIGLAEG LESEKKAIVP 780
 LVIVSALTPI CLVVLVGLI YWRKCPQTAH FYLEDSTSPR VISTPPTPIF PISDDVGAIP 840
 IKHFPKHVAD LHASSGTFEE FETLKEFYQE VQSCVDLGI TADSSNHPDN KHKRYINIV 900
 AYDHSRKLVA QLAEKDGKLT DYINANYVDG YNRKAYIAA QGPLKSTAE FWRMIWEHN 960
 30 EVVMITNLV EKRRTKCDQY WPAAGSEYQ NPLVTQKSVQ VLAYYTVRNF VHCASGVGRT 1020
 SQKGRPSGRV VQYHYHTQMP DMGVPEYSLP VLTFRVKAAY AKRHAVGPVV VHCASGVGRT 1080
 GTYIVLDSML QOIQHEGTVM IFGLKHRS QRNYLVQTEB QYVFIHDLTV EAILSKEV 1140
 LDSHIHAYVN ALLIPGPAK TKLEKQFQGL TSPRLCEGR TISAHCNPLP PGLTDPPTSA 1200
 SRVAGTILLS QSNIQSDYS AALKQCNREK NRTSSIIPE RSRVGISLS GEGTDYINAS 1260
 35 YIMGYQSNL FIIOHLLH TIKDFWRMIW DHNAQLVMI PDGQMAEDE FVYWPNDKEP 1320
 INCESFKVTL MAEHEKCLSM EEKLIQDFI LEATQDDYVL EVRHFQCFKW PNPDSPIST 1380
 FELISVKEE AANRDGPMIV HDEHGGVITAG TFCALTILMH QLEKENSVDV YQVAMINLM 1440
 RGVFADIEQ YQFLYKVLIS LVGTRQENP STSLDSNGAA LPDGNIAESL ESLV

Seq ID NO: 582 DNA sequence
 Nucleic Acid Accession #: NM_002851.1
 Coding sequence: 148..7092

1 11 21 31 41 51
 CACACATACG CACGCACGAT CTCACCTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60
 CAAAAAAGAC ATTTCCCTCG CTCCCTCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120
 CGGCGAGGCG CGCGACAGCG TCTGGAATG CGAATCTCAA AGCGTTCTCT CGCTTGCAAT 180
 50 CAGCTCTCTT GTGTTTGGCG CTCTGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
 CTGTGTGAAG AGATTGGCTG GTCCATATCA GGAGCACTGA ATCAAAAAA TTGGGGAAAG 300
 AAATATCCAA CATGTAATAG CCCAAAAA TCTCTATCA ATATTGATGA AGATCTTACA 360
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTGAGGGTT GGGATAAAAC ATCATTGGAA 420
 AACACATTCA TTCATAACAC TGGGAAAAA GTGGAATTA ATCTCACTAA TGACTACCGT 480
 55 GTCACGCGAG GAGTTTCAGA AATGGTGT TTAAAGCAAGCA AGATAACTT TCACTGGGGA 540
 AAATGCAATA TGTCATCTGA TGATCGAGAG CATAGTTTAG AAGGACAAA ATTTCCACTT 600
 GAGATGCAAA TCTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAA 660
 GGAAAGGGA AGTTAAGAGC TTTATCCATT TTGTTGAGG TTGGGACAGA AGAAAAATTG 720
 GATTTCAGAG CGATTATTGA TGGAGTCGAA AGTGTGTAGT GTTTTGAGGA CGAGGCTGCT 780
 60 TTGATCTCAT TCATCTGTT GAACCTCTG CCAACTCAA CTGACAAGTA TTACATTAC 840
 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGATTGT TTTTAAAGAT 900
 ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTTGTG AAGTCTTAC AATGCAACA 960
 TCTGTTATG TCATGCTGAT GGACTACTTA CAAAAAATT TTGAGAGCA ACAGTACAAG 1020
 TTCTCTAGAC AGGTGTTTTT CTCATACACT GGAAGGGAAG AGATTCTAGA AGCAGTTTGT 1080
 65 AGTTGAGGAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCTT TCTTGTACA 1140
 TGGGAAAGAC CTCAGTCTGT TATGATACC ATGATTGAGA AGTTTGAGT TTTGTACCAG 1200
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260
 GGTGCTATT TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320
 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTCGACAT GCCTACTGAT 1380
 70 AATCTGAAC TTGATCTTTT CCTGAATTA ATTGGAACGT AAGAAATAAT CAAGGAGGAG 1440
 GAAGAGGGAA AAGACATTGA AGAAGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500
 AACCAATCA GGAAGGGAAG ACCCCAGATT TCTACCACAA CACACTACAA TGCATAGGG 1560
 ACGAATACA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620
 AAGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680
 75 ACAGAAAAG ATATTTCCTT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740
 GAAGGTACTT CAGCTCTTTT AAATGATGGC TCTAAACTGT TTCTTAGATC TCCACATATG 1800
 AACTTGTGGG GAGTGCAGA ATCTTAAAT ACAGTTTCTA TAACAGAAATA TGAGGAGGAG 1860
 AGTTTATTGA CCAATTTCAG GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCAGTCCC 1920
 GCAACTCTG CTATCCCATC CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980
 80 GAAACCCAG AGACAATAAC ATATGATGTC CTATACCAG AATCTCTAG AAATGCTTCC 2040
 GAAGATTCAA CTTATCAGG TTCAGAGAA TCACTAAAG ATCTTCTAT GGAGGGAAAT 2100
 GTGTGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160
 AGCTTTCTCC AGACTAATTA CACTGAGATA CGTGTGATG AATCTGAGAA GACAACCAAG 2220
 TCCTTTCTG CAGGCCAGT GATGTACAG GGTCCCTCAG TTACAGATCT GGAATGCCA 2280
 85 CATTATCTA CTTTGCCTA CTCCCAACT GAGTAACAC CTCACTCTT TACCCATCC 2340
 TCCAGACAC AGGATTGTGT CTCACGGTC AACGTGGTAT ACTGCAGAC AACCCAACCG 2400
 GTATACAATG GTGAGACACC TCTTCAACT TCTACAGTA GTGAAGTCTT TCCTCTAGTC 2460
 ACCCTTTGTG TGCTTGACAA TCAGATCCTC AACACTACCC CTGCTGCTC AAGTAGTGAT 2520

	TCGGCCTTGC	ATGCTAGGCC	TGTATTTCCT	AGTGTGATG	TGTCATTGA	ATCCATCCTG	2580
	TCTTCCCTATG	ATGTTGACCC	TTTGCTTCCA	TTTTCCTCTG	CTTCTCTCAG	TAGTGAATTG	2640
	TTTCGCCATC	TGCATACAGT	TTTCAAAATC	CTTCCACAAG	TTACTTTCAG	TACCGAGAGT	2700
5	GATAAGGTGC	CCITGTCATG	TTCTCTGCCA	GTGGCTGGGG	GTGATTGCT	ATTAGAGCCC	2760
	AGCCTTGCTC	AGTATTCTGA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GACGCTGGAA	2820
	TTTGSTAGTG	AATCTGGTGT	TCTTTATAAA	ACGCTTATGT	TTCTCAAGT	TGAACCAACC	2880
	AGCAGTGTATG	CCATGATGCA	TGCACSTTCT	TCAGGGCCTG	AACCTTCTTA	TGCTTGTCT	2940
	GATAATGAGG	GCTCCCAACA	CATCTTCACT	GTTCCTTACA	GTTCGTGCAAT	ACCTGTGCAT	3000
10	GATTCTGTGG	GTGTAACCTA	TCAGGGTTCC	TTATTTAGCG	GCCTTAGCCA	TATACCAATA	3060
	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGATGGGG	AATGGTCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAAACAT	TTCTTCACTT	GTTCCTGTAG	CTGAATTTC	ATATACAACA	3240
	TCTGTGTTTG	GTGATGATAA	TAAAGGCGCT	TCTAAAAGTG	AAATAATATA	TGGAATGAG	3300
	ACTGAACCTG	AAATTCTCTC	TTTCAATGAG	ATGGTTTACC	CTTCTGAAAG	CACAGTCATG	3360
15	CCCAACATGT	ATGATAATGT	AAATAAGTTG	AATGCGTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTCTTAGCA	CCAAGGGCAT	GTTCCTGCTC	TCCCTTGCTC	ATACCAACC	TAAAGTTTCT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAAT	AACCTTTTCAG	TTCAACCTTAC	ACATACTGTC	3540
	TCTCAAGCAT	CTGCTGACAC	TTGCTTTAAA	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
	TCCTCTGACC	CTGCTTCTAG	TGAAATGTTA	TCTCCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
20	ACCTCAGCTT	CTTTTAGTAC	TGAAGTATTG	CTACAACCTT	CCTTTTCAGGC	TTCTGATGTT	3720
	GACACCTTGC	TAAAACCTGT	TCTTCCAGCT	GTGCCAGTGT	ATCCAATATT	GGTTGAAACC	3780
	CCCAAGTTTG	ATAAAATTAG	TTCTACAATG	TTGCATCTCA	TTGTATCAAA	TTCTGCTTCA	3840
	AGTGAAACAA	TGCTGCACTC	TACATCTGTA	CCAGTTTITG	ATGTGTGCGC	TACTTCTCAT	3900
25	ATGCACCTCT	CTTCACTTCA	AGGTTTGACC	ATTTCTCTATG	CAAGTGAGAA	ATATGAACCA	3960
	GTTTGTGTTA	AAAGTGAAAG	TTCCCAACAA	GTGGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCAAA	CGGCCAATTT	GGAGATTAAC	CAGGCCATC	CCCCAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTATTC	AATTGATGAA	CCATTAAATA	CACATAATAA	TAAAGTTTATA	4140
	CATTCCGATG	AAATTTTAA	CTCCACCAAA	AGTTCTGTGA	CTGGTAAGST	ATTTGCTGGT	4200
30	ATTCCACAG	TTGCTTCTGA	TACATTTGTA	TCTACTGATC	ATTCTGTTC	TATAGGAAAT	4260
	GGGCATGTTG	CCATTACAGC	TGTTTCTCCC	CACAGAGATG	GTTCGTAAAC	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACCTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCGGT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGACAGA	4440
	GATAGTGATG	GCTTATCCAT	TCCATAAGTG	ATGTCATGCT	CATCCTATAG	AGAATCACAG	4500
35	GAAAAGGTAA	TGAATGATTC	AGACACCCAC	GAAAACAGTC	TTATGGATCA	GAATAATCCA	4560
	ATCTCATACT	CACATCTGTA	GAATCTGAA	GAAGATAATA	GAGTCACAAG	TGATCTCTCA	4620
	GACAGTCAAA	CTGGTATGGA	CAGAAGTCTT	GGTAAATCAC	CATCAGCAAA	TGGGCTATCC	4680
	CAAAAGCACA	ATGATGGAAA	AGAGGAAAAT	GACATTGAGA	CTGGTAGTGC	TCTGCTTCTC	4740
	CTCAGCCCTG	AATCTAAAGC	ATGGGCAGTT	CTGACAAGTG	ATGAAGAAAG	TGGATCAGGG	4800
40	CAAGGTACCT	CAGATAGCCT	TAATGAGAA	GAGACTTCCA	CAGATTTCAG	TTTTGTCAGC	4860
	ACTAATGAAA	AAGATGCTGA	TGGGATCCTG	GCAGCAGGTG	ACTCAGAAAT	AACTCCTGGA	4920
	TTCCACAGT	CCCCAACATC	ATCTGTTACT	AGCGAGAACT	CAGAAGTGT	CCACGTTTCA	4980
	GAGGCAGAGG	CCAGTAATAG	TAGCCATGAG	TCTCGTATTG	GTCTAGCTGA	GGGGTTGGAA	5040
	TCCGAGAAGA	AGGCAGTTAT	ACCCCTTGTG	ATCGTGTGAG	CCCTGACTTT	TATCTGTCTA	5100
45	GTGGTCTCTG	TGGGTATTCT	CATCTACTGG	AGGAAATGCT	TCCAGACTGC	ACACTTTTAC	5160
	TTAGAGGACA	GTACATCCCC	TAGAGTTATA	TCCACACCTC	CAACACCTAT	CTTTCCAATT	5220
	TCAGATGATG	TCCGAGCAAT	TCCAATAAAG	CACCTTCCAA	AGCATGTTGC	AGATTTCAT	5280
	GCAAGTAGTG	GGTTTACTGA	AGAATTTGAG	ACACTGAAAG	AGTTTACCAC	GGAAAGTGCAG	5340
	AGCTGTACTG	TTGACTTAGG	TATTACAGCA	GACAGCTCCA	ACCACCCAGA	CAACAAGCAC	5400
50	AAGAATCGAT	ACATAAATAT	CGTTGCCTAT	GATCATAGCA	GGGTTAAGCT	AGCACAGCTT	5460
	GCTGAAAAGG	ATGGCAAACT	GACTGATTAT	ATCAATGCCA	ATTATGTTGA	TGGGTACAA	5520
	AGACCAAAAG	CTTATATTGC	TGCCCAAGGC	CCACTGAAAT	CCACAGCTGA	AGATTCTTGG	5580
	AGAATGATAT	GGGAACATAA	TGTGGAAGTT	ATTGTCTATG	TAAACAACTT	CGTGAGAGAA	5640
	GGAAAGGAAA	AATGTGATCA	GTACTGGCCT	GCAGTGGGA	GTGAGGAGTA	CGGGAACCTT	5700
55	CTGGTCACTG	AGAAGAGTGT	GCAAGTGCTT	GCCTATTATA	CTGTGAGGAA	TTTTACTCTA	5760
	AGAAACACAA	AAATAAAAAA	GGGCTCCAG	AAAGGAAGAC	CCAGTGGACG	TGTGTACACA	5820
	CAGTATCACT	ACACCGCAGT	CTATGCCAAG	CGCCATGCG	TGGGCGCTGT	TGTGTCAC	5880
	ACCTTTGTGA	GAAAGGCAGC	CTATGCCAAG	CGCCATGCG	TGGGCGCTGT	TGTGTCAC	5940
	TGCACTGCTG	GAGTTGGAAG	AACAGGCACA	TATATTGTGC	TAGACAGTAT	GTTCAGCAG	6000
60	ATTCAACACG	AAGGAAGTGT	CAACATATTT	GGCTTCTTAA	AACACATCCG	TTTCAAAAGA	6060
	AATTATTGTT	TACAACTGTA	GGAGCAATAT	GTCTTCTATC	ATGATACACT	GGTTGAGGCC	6120
	ATACTTAGTA	AAGAACTGTA	GGTGCTGGAC	AGTCATATTC	ATGCCATATG	TAATGCACTC	6180
	CTCATTTCTG	GACCAAGCAG	CAAAACAAAG	CTAGAGAAAC	AATTCAGACT	CCTGAGCCAG	6240
	TCAATATATC	AGCAGAGTGA	CTATTCTGCA	GCCTTAAAGC	AATGCAACAG	GGAAAGAAAT	6300
65	CGAACTTCTT	CTATCATCCC	TGTGGAAAGA	TCAAGGGTTG	GCATTTCATC	CCTGAGTGA	6360
	GAAGGCACAG	ACTACATCAA	TGCTTCTTAT	ATCATGGGCT	ATTACCCAG	CAATGAATTC	6420
	ATCATTTCCC	AGCACCTCTT	CCTTCATACC	ATCAAGGATT	TCTGGAGGAT	GATATGGGAC	6480
	CATAATGCCC	AATCTGGTGT	TATGATTCTT	GATGGCCAAA	ACATGGCAGA	AGATGAATTT	6540
	GTTTACTGGC	CAATAAAGA	TGAGCCTATA	AATTGTGAGA	GCTTTAAGGT	CACCTTTATG	6600
70	GCTGAAGAAC	ACAAATGTCT	ATCTAATGAG	GAAAACTTAA	TAATTACAGTA	CTTTATCTTA	6660
	GAAGCTACAC	AGGATGATTA	TGTACTTGAA	GTGAGGCACT	TTCAAGTPTC	TAAATGGCCA	6720
	AATCCAGATA	GCCCCATTAG	TAAAACTTTT	GAACTTATAA	GTGTTATATA	AGAAGAAGCT	6780
	GCCAATAGGG	ATGGGCTTAT	GATTGTTTAT	GATGAGCATG	GAGGAGTGAC	GGCAGGAATC	6840
	TTCTGTGCTC	TGACAAACCTT	TATGACCAAA	CTAGAAAAG	AAAATTCCTG	GGATGTTTAC	6900
75	CAGGTAGCCA	AGATGATCAA	TCTGATGAGG	CCAGGAGTCT	TTGCTGACAT	TGAGCAGTAT	6960
	CAGTTTCTCT	ACAAAGTGTAT	CCTCAGCCTT	GTGAGCACAA	GGCAGGAAGA	GAATCCATCC	7020
	ACCTCTCTGG	ACAGTAATGG	TGCAGCAATT	CCTGATGGAA	ATATAGCTGA	GAGCTTAGAG	7080
	TCCTTAGTGT	AACACAGAAA	GAGTCAACAT	TGAGCATGTT	TTTCTCTTTC	ATCACTGTAC	7140
	CTAAAATTAG	CGAGGAAAT	CAGTCTAGTT	CTGTTATCTG	TTGATTTCCT	ATCACTGTAC	7200
	AGTAACTTTC	ATGACATAGG	ATTCTGCGGC	CAAAATTATA	TCATTAAACA	TGTGTGCTTC	7260
80	TTTGCAAGAC	TTGTAATTTA	CTTATTATGT	TTGAACATAA	ATGATTGAAT	TTTACAGTAT	7320
	TTCTAAGAA	GGAAATGTGG	TATTTTTTTC	TGATATTGAT	TTAACAGAAA	ATTTCAATTT	7380
	ATAGAGGTTA	GGAAATTCAC	ACTACAGAAA	ATGTTTGTIT	TTAGTGTCAA	ATTTTATGCT	7440
	GTATTGTGTC	CAATTATCAG	GTTTGTGAGA	AAATATAACT	TTAATACAGT	AGCCTGTAAA	7500
	TAAACACATC	TTCCATATGA	TATTCACAT	TTTACAACCT	CAGTATTAC	CTAAAGTAGA	7560
85	ATAATCTGTT	TACTTATTGT	AAATACTGCC	CTAGTGTCTC	CATGAGCCAA	ATTTATATTT	7620
	ATAATTTAG	ATTTTATAT	TTTACTACTG	AGTCAAGTTT	TCTAGTTCTG	TGTAATTTGT	7680
	TAGTTAATG	ACGTAGTTCA	TTAGCTGTGC	TTACTCTACC	AGTTTCTGTA	CATTGTATTG	7740

TGTTACCTAA GTCAATTAAC TTGTTTCAGC ATGTAATTTT AACTTTTGTG GAAATAGAA 7800
 ATACCTTCAT TTGAAAGAA GTTTTATGA GAATAACACC TTACCAACA TTGTTCAAAT 7860
 GGTTTTTATC CAAGGAATG CAAAAATAA TATAAATATT GCCATTAAAA AAAAAAATAA 7920
 AAAAAAAAAA AAAAAAAAAA A

Seq ID NO: 583 Protein sequence
 Protein Accession #: NP_002842.1

1 11 21 31 41 51
 10 MRILKRFLAC IQLLCVCRLD WANGYYRQR KLVVEIGWSY TGALNQKNWG KKYPTCNSPK 60
 QSPINIDEDL TQVNVNLKLL KFGQWDXSL ENTFTHNTGK TVSINLTNDY RVSGGVSEMV 120
 FKASKITFWH GKCMSSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEAV KGRKGLRALS 180
 ILFEVGTEN LDFKAIIDGV ESVSFRGKQA ALDPFILLNL LPNSTDKYI: YNGSLTSPPC 240
 15 TDTVDWIVFK DTVSISESQL AVFCEVLTMO QSGYVMLMDY LQNNFREQQY KFSRQVFSY 300
 TGKEBIHEAV CSSEPEHVQA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360
 HEFLTGDYQD LGAILNLLP NMSYVLQIVA ICTNGLYKGY SDQLIVDMPT DNPBLDLFPE 420
 LGTEELIKE EEEGKIDEG AIVNPGRDSA TNQIRKKEPO ISTTHYNRI GTKYNEAKTN 480
 RSPTRGSEPS GKGDPVNTSL NSTSQPVTKL ATEKDISLTS QTVTELPHT: VEGTSASLND 540
 20 GSKTVLRSPL MNLSGTAESL NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600
 ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDPSEMG NVWFPSSTDI 660
 TAQPDVGSGR ESFLQNTYTE IRVDESEKTT KSFSAGPVMS QGFSVTDLEM PHYSTFAYFP 720
 TEVTPHAFT SSROQDLVST VNVVYSQTQ PVNGETPLQ PSYSSEVPPL VTPLLLDNQI 780
 LNTTPAASSS DSALHATPVF PSVDVSFESI LSSYDGAPLL PFSSASFSE LFRHLHTVSQ 840
 25 ILPQVTSATE SDKVLHASL PVAGGDLLE PSLAQYSDVL STTHAASETL EFGSESGVLY 900
 KTLMFQVEP PSSDMMHAR SSGPEPSYAL SDNEGSQHIF TVSYSSAIPV HDSVGVTYQG 960
 SLFSGSPHIP IPKSSLTITP ASLLQPTHAL SGDGEWSGAS SDSEFLPDT DGLTALNISS 1020
 PVSVAEFTYT ESFVGDNDKA LSKSEIYGN: ETELQIPSFN EMVPSSESTV MPNMVDNVNK 1080
 LNASLQETSV SISSTKGMFP GSLAHTTKV FDHEISQVPE NNFVQPTHT: VSQASGDTSL 1140
 30 KPVLSANSEP ASSDPASSEM LSPSTQLLFY ETSASFSTEV LLQPSFQASD VDTLLKTULP 1200
 AVPSDPILVE TPVKDKISST MLHLIVNSA SSENMLHSTS VPFVDVSPTS HMHSASLQGL 1260
 TISYASKEYE PVLLKSESSH QVPSLYSND ELFQTNALIE NQAHPPKGRH VFATPVLSID 1320
 EPLNTLNLKL IHSDEILTST KSSVTGKVEA GIPTVASDTF VSTDHSPVIG NGHVAITAVS 1380
 PHRDGSVTST KLLFPKATS: ELSHSAKSDA GLVGGGEDGD TDDGDDDDDD RDSGLSIHK 1440
 35 CMSCSSYRES QEKVMDSDT HENSLMDQNN PISYSLSENS EEDNRVTSVS SDSQTMDRS 1500
 PGKSPSANGS SQKHNDGKEE NDIQTGSALL PLSPEKAWA VLTSDEESGS GQGTSDSLNE 1560
 NETSTDFSA DTNEKDADGI LAAGDSEITP GFQPSSTSSV TSENSEVPHV SEAEASNSSH 1620
 ESRIGLABGL ESEKKAIVPL VIVSALTFC LVVLVGLIY WRKCFQTAHF YLEDSTSPRV 1680
 40 ISTPTPIFP ISDDVGAIP KHFPKHVADL HASSGTFEEF ETLKEFYQEV QSCSTDGLIT 1740
 ADSSNNHNDK HKNRYINIVA YDHSRVKLAQ LAEKDGKLT D YINANYVDGY NRPKAYIAAQ 1800
 GPLKSTAEF WRMWHEHNV VIVMITNLVE KGRKCDQYW PADGSEEGN FLVTKQSVQV 1860
 LAYTVRNTF LENTKIKKGS QKGRPSGRVV TQVHYTQWPD MGVPYSLPV LTFVRKAAVA 1920
 KRHAVGPVV HCSAGVGRGT TYIVLDSMLQ QIQHEGTVNI FGLKHRSQ RNYLVQTEEQ 1980
 YVFIHDLVE AILSKETEVL DSHIHAYVNA LLIPGAGTK KLEKQFQLLS QSNIIQSDYS 2040
 45 AALKQCNREK NRTSSIPVE RSRVGISSLS GEGTDYINAS YIMGYQSNF FIITQHPLH 2100
 TIKDFWRMIW DHNAQLVVM: PDGQNMABE FVWPNKDEP INCESPKVTL MAEHHKCLSN 2160
 EEKLIQDFI LEATQDDYVL EVRHFQCPKM PNPDSPIKST FELISVIEE ANRDGPMIV 2220
 HDEHGGVTA TFCALTLMH QLEKENSVDV YQVAKMINLM RPFVFADIEQ YQFLYKVLIS 2280
 LVSTRQENP STSLDSNGAA LPDGNIAESL ESLV

Seq ID NO: 584 DNA sequence
 Nucleic Acid Accession #: NM_005688.1
 Coding sequence: 126..4439

1 11 21 31 41 51
 55 CCGGCGAGGT GGCTCATGCT CCGGAGCGTG GTTGAGCGGC TGCGCGGTT GTCTGGAGC 60
 AGGGCGCGAG GAATTCGTGAT GTGAACTAA CAGTCTGTGA GCCCTGGAAC CTCCTCAGC 120
 60 AGAAGATGAA GGATATCCAG ATAGGAAAG AGTATATCAT CCCAGTCTCT GGTATAGAA 180
 GTGTGAGGGA GAGAACAGC ACTTCGGGA CGCAGAGAGA CCGTGAAGAT TCCAAGTTCA 240
 GGAGAACTCG ACOGTGGAA TGCCAAGATG CCTTGGAAAC AGCAGCCCGA GCCGAGGGCC 300
 TCTCTCTTGA TGCTCCATG CATTCTCAGC TCAGAACTCT GGATGAGGAG CATCCCAAGG 360
 GAAAGTACCA CATGTGCTG AGTGCTCTGA AGCCCATCCG GACTACTTCC AAACACCCAGC 420
 65 ACCCAGTGGG CAATGCTGGG CTTTTCCTCT GTATGACTTT TCGTGGCTT TCTCTCTGG 480
 CCCGTGTGGC CCACAAGAAG GGGGAGCTCT CAATGGAAGA CGTGTGCTCT CTGTCCAAGC 540
 ACGAGTCTTC TGACGTGAAC TGCAGAAGAC TAGAGAGACT GTGGCAAGAA GAGCTGAATG 600
 AAGTGTGGCC AGACGCTGCT TCCCTGCGAA GGGTGTGTG GATCTTCTGC CGCACCAGGC 660
 TCATCTGCTC CATCGTGTGC CTGATGATCA CGCAGCTGGC TGGCTTCAGT GGACAGCCT 720
 70 TCATGGTGAA ACACCTCTTG GAGTATACCC AGGCAACAGA GTCTAACCTC CAGTACAGCT 780
 TGTGTGAGT GCTGGGCTCT CTCCTGAGCG AAATCGTGC GTCTTGGTCG CTGCACTGA 840
 CTGGGGCATT GAATACCGA ACCGTGTGCC GCTTGGGGG GCCATCCTA ACCATGGCAT 900
 TTAAGAGATC CCTTAAGTTA AAGAACATTA AAGAGAAATC CCTGGGTGAG CTCATCAACA 960
 75 TTTGCTCCAA CGATGGGCAG AGAATGTTG AGGCAGCAGC CGTGGCAGC CTGCTGGCTG 1020
 GAGGACCCGT TGTGCGCAT TTAGGCATGA TTTATAATGT AATTATTCTG GGACCAACAG 1080
 GCTTCTCTGG ATCAGCTGTT TTTATCCTCT TTTACCCAGC AATGATGTTT GCATCAAGC 1140
 TCACAGCATA TTTAGGAGA AAATGCGTGG CGGCCACGGA TGAACGTGTC CAGAAGATGA 1200
 ATGAAGTTCT TACTACATT AAATTTATCA AAATGTATGC CTGGGTCAA GCATTTTCTC 1260
 AGAGTGTTC AAAAAATCCG GAGGAGGAGC GTCGGATATT GAAAAAGCC GGTACTTCC 1320
 80 AGGTATCAC TGTGGGTGTG GCTCCCATG TGGTGTGAT TGCCAGCGTG GTGACCTCT 1380
 GTGTTTCAT GACCTGGGC TTCGATCTGA CAGCAGCACA GCTTTCACA GTGTGACAG 1440
 TCTTCAATC CATGACTTTT GCTTGAAG TAACACCGTT TTAGTAAAG TCCTCTCAG 1500
 AAGCCTCAGT GCTGTTGAC AGATTAAAGA GTTGTGTTCT AATGGAAGAG GTTCACATGA 1560
 TAAAGAACAA ACCAGCCAGT CCTCACATCA AGATAGAGAT GAAAAATGCC ACCTGGCAT 1620
 85 GGGACTCCTC CCACTCCAGT ATCCAGAACT CGCCCAAGCT GACCCCAAAA ATGAAAAAG 1680
 ACAAGAGGGC TTCCAGGGGC AAGAAAGAGA AGGTAGAGCA GCTGCAGCGC ACTGAGCATC 1740
 AGCGGTGCT GCACAGCAG AAAGGCCACC TCCTCTGGA CAGTAGCAGG CGGCCAGTTC 1800
 CCGAAGAGGA AGAAGCCAG CACATCCACC TGGGCCACCT GCGCTTACAG AGGACACTGC 1860

ACAGCATCGA TCTGGAGATC CAAGAGGGTA AACTGGTTGG AATCTGCGGC AGTGTGGGAA 1920
 GTGGAAAAAC CTCTCTCAAT TCAGCCATTT TAGGCCAGAT GACGCTTCTA GAGGGCAGCA 1980
 TTGCAATCAG TGGAACTTTC GCTTATGTGG CCCAGCAGGC CTGGATCCTC AATGCTACTC 2040
 TGAGAGACAA CATCTCTGTT GGAAGGAAT ATGATGAAGA AAGATACAAC TCTGTGCTGA 2100
 ACAGCTGCTG CCTGAGGCTT GACCTGGCCA TTCTTCCAG CAGCGACCTG ACGGAGATTG 2160
 GAGAGCGAGG AGCCAACTG AGCGTGGGC AGCGCCAGAG GATCAGCCTT GCCCGGGCTG 2220
 TGTATAGTGA CAGGAGCATC TACATCTGGG ACGACCCCTT CAGTGCCCTA GATGCCCATG 2280
 TGGGCAACCA CATCTTCAAT AGTGTATCC GGAACATCT CAAGTCCAAG ACAGTTCTGT 2340
 TTGTTACCCA CAGTTTACAG TACCTGGTTG ACTGTGATGA AGTGTCTTC ATGAAAGAGG 2400
 GCTGTATTAC GGAAGAGGC ACCATGAGG AACTGATGAA TTTAAATGGT GACTATGCTA 2460
 CCATTTTAA TAACCTGTTG CTGGGAGAGA CACCGCCAGT TGAGATCAAT TCAAAAAAGG 2520
 AAACCACTGG TTCACAGAAG AAGTCACAAG ACAAGGGTCC TAAACACAGA TCAGTAAAGA 2580
 AGGAAAAAGC AGTAAAGCCA GAGGAAGGGC AGCTTGTGCA GCTGGAAGAG AAAGGGCAGG 2640
 GTTCAGTGCC CTGGTCAGTA TATGGTGCTT ACATCCAGGC TGCTGGGGGC CCTTGGGCAT 2700
 TCTGCTTAT TAATGGCCCTT TCATGTCTGA ATGTAGGCAG CACCGCCTTC AGCAGCTGGT 2760
 GGTGAGGTTA CTGGATCAAG CAGGAAGCG GGAACACCA TGTGACTCGA GGAACAGAGA 2820
 CCTCGGTGAG TGACAGCATG AAGGACAATC CTCATATGCA GTACTATGCC AGCATCTACG 2880
 CCTCTCCAT GGCAGTCATG CTGATCCTGA AAGCCATTCG AGGAGTTGTC TTTGTCAGG 2940
 GCAGCGTGG AGCTTCTCTC CGGCTGCATG ACGAGCTTTT CCGAAGGATC CTTTGAAGCC 3000
 CTATGAAGTT TTTTGACACG ACCCCACAG GAGGATTCT CAACAGGTTT TCCAAAGACA 3060
 TGGATGAAGT TGACGTGGG CTGCGTTC AGGCGAGAT GTTCATCCAG AACGTTATCC 3120
 TGGTGTCTT CTGTGTGGGA ATGATCGCAG GAGTCTTCCC GTGGTTCTT GTGGCAGTGG 3180
 GGCCCTTGT CATCTCTT TCAGTCTCTC ACATTGTCTC CAGGGTCTCT ATTCCGGAGC 3240
 TGAAGCGTCT GGACAATATC ACGCAGTCAC CTTCCTCTC CCACATCAC TCCAGCATAC 3300
 AGGGCTTGC CACCATCCAC GCCTACAATA AAGGGCAGGA GTTCTGCAC AGATACCAAG 3360
 AGCTGCTGGA TGACAAGCAA GCTCTTTT TTTGTTTAC GTGTGCGATG CGGTGGCTGG 3420
 CTGTGCGGCT GGACCTCATC AGCATCGCCC TCATCACCA CACGGGCTG ATGATCGTTC 3480
 TTATGACGCG GCAGATTCCC CCAGCCTATG CCGGTCTCGC CATCTCTTAT GCTGTCCAGT 3540
 TAAAGGGGCT GTTCAAGTTT ACGTCAGAG TGGCATCTGA GACAGAAGCT CGATTACCT 3600
 CCGTGGAGAG GATCAATCAC TACATTAAGA CTCTGTCTT GGAAGCACTT GCCAGAATTA 3660
 AGAACAGAGG TCCCTCCCTT GACTGGCCCC AGGAGGGAGA GGTGACCTTT GAGAAGCAG 3720
 AGATGAGGTA CCGAGAAAAC CTCCTCTCTG TCCTAAAGAA AGTATCCTTC ACGATCAAA 3780
 CTAAGAGAAA GATTGGCATT GTGGGGCGGA CAGGATCAGG GAAGTCTCTG CTGGGGATGG 3840
 CCCTCTTCCG TCTGGTGGAG TTATCTGGAG GCTGCATCAA GATTGATGGA GTGAGAATCA 3900
 GTGATATTGG CCTTGGCAGC CTCCGAAGCA AACTCTCTAT CATCTCTCAA GAGCCGGTGC 3960
 TGTTCAGTGG CACTGTGAGA TCAATTGG ACCCTTCAA CCACTACACT GAAGACAGGA 4020
 TTTGGGATGC CCGTGGAGAG ACACACATGA AAGAATGTAT TGCTCAGCTA CCTCTGAAAC 4080
 TTGAATCTGA AGTGTGGAG AATGGGGATA ACTTCTCAGT GGGGGAACGG CAGCTCTTGT 4140
 GCATAGCTAG AGCCCTGCTC CGCCACTGTA AGATTCTGAT TTTAGATGAA GCCACAGCTG 4200
 CCATGGACAG AGAGACAGAG TTATTGATTC AAGAGACCAT CCGAGAAGCA TTTGCAGACT 4260
 GTACCATGCT GACCATTGGC CATCGCCTGC ACACGGTTCT AGGCTCCGAT AGGATTATGG 4320
 TGCTGGCCCA GGGACAGGTG GTGAGTTTG ACACCCATC GGTCTTCTG TCCAACGACA 4380
 GTTCCCGATT CTATGCCATG TTGTCTGCTG CAGAGAACAA GGTCTCTGTC AAGGGCTGAC 4440
 TCCTCCCTGT TGACGAAGTC TCTTTCTT AGAGCATTGC CATCTCCCTG CTGGGGCGGG 4500
 CCCCTCATCG CGTCTCTCTA CCGAAACCTT GCCTTCTCG ATTTTATCTT TCGCACAGCA 4560
 GTTCCGAGT GCTTGTGTG TTTCACTTT AGGAGAGTCT ATATTTTGAT TATTGTATT 4620
 ATTCCATATT CATGTAAACA AATTTAGTT TTTGTTCTTA ATTGCACTCT AAAAGGTTCA 4680
 GGAACCGGTT ATTATAATTG TATCAGAGGC CTATAATGAA GCTTTATACG TGTAGCTATA 4740
 TCTATATATA ATTCTGTACA TAGCCTATAT TTACAGTGAA AATGTAAGCT GTTTATTTTA 4800
 TATTAATAAT AGCATGTGCG TAATAACAGT GCATATTCTT TCTATCATT TTTGTACAGT 4860
 TTGCTGTACT AGAGATCTGG TTTTGTCTAT AGACTGTAGG AAGAGTAGCA TTTCTTCTT 4920
 CTCTAGCTGG TGGTTTACAG GTGCCAGGTT TCTGSGTGT CCAAGGAAG ACGTGTGGCA 4980
 ATAGTGGGCC CTCCGACAGC CCCCTCTGCC GCCTCCCCAC AGCCGCTCCA GGGGTGGCTG 5040
 GAGACGGGTG GGCGCTGGA GACCATGCG AGCGCGTGA GTTCTCAGG CTCTGCTCTT 5100
 CTGTCTGTGT GTCACTTACT GTTCTGTCA GAGAGCAGC GGGGCGAAGC CAGGCGCCCT 5160
 TTTCACTCCC TCCATTAAGA ATGGGATCA CAGAGACATT CCTCCGAGCC GGGGAGTTTC 5220
 TTTCTGCTCT TCTCTTTT GCTGTGTTT CTAACAAGA ATCAGTCTAT CCACAGAGAG 5280
 TCCCCTGCC TCAGGTTCTT ATGGCTGGCC ACTGCACAGA GCTCTCCAGT TCCAGACCT 5340
 GTTGGTTTCCA AGCCCTGAG CCAACTGCTG CTTTTGGAG TGGCACTTTT TCATTGCTCT 5400
 ATTCCACAG CTCCACAGTT CAGTGGCAGG GCTCAGGATT TCGTGGGTCT GTTTCTCTT 5460
 CTCACCGCAG TGTGCGACA GTCTCTCTCT CTCTCTCCCC TCAAGTCTG CAACCTTAAG 5520
 CAGCTCTGCT TAATCAGTGT CTCACACTGG CGTAGAAGTT TTTGTACTGT AAAGAGACCT 5580
 ACCTCAGGTT GCTGGTGTCT GTGTGGTTTG GTGTGTCTCC GCAAAACCCC TTTGTGCTGT 5640
 GGGGCTGGTA GCTCAGGTGG GCGTGTCTAC TGCTGTCTAC AGTTGAATGG TCAGGTTGCT 5700
 ATGTCGTGAC CAATAGACA TTCTGTGCC TTAGCATGTT TGCTGAACAC CTTGTGGAAG 5760
 AAAAAATCTG AAAATGTGAA TAAATATTATT TTGATTTTTG TAAAAAAA AAAAAAAA 5820
 AAAAAAAA AAAAAA

Seq ID NO: 585 Protein sequence
 Protein Accession #: NP_005679.1

1 11 21 31 41 51
 MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFR TRPLECQDAL ETAARAEGLS 60
 LDASHMSQLR ILDEBHPKXK YHGLSALKP IRTTSKHQHP VDNAGLFSCM TFSWLSLLAR 120
 VAHKKGELSM EDVWSLSKHE SSDVNCRRLE RLWQBELNEV GPDAASLRV VWIFCRTRLI 180
 LSIVCLMITQ LAGFSGPAPM VKHLEYTQA TESNLQYSL LVLGLLLTEI VRSWSLALTW 240
 ALNYRTGVRL RGAILTMAFK KILKLNKIKE KSLGELINIC SNGQRMFEA AAVGSLLAGG 300
 PVVAILGMIY NVIIIGPTGF LGSVAFILFY PAMMFASRLT AYFRRCVAA TDERVQKQNE 360
 VLTVIKFKIM YAWVKAFQSQ VQKIREEERR ILEKAGYFQG ITVGVAIVV VIASVVTFSV 420
 HMTLGFDLTA AQATVTVTVF NSMTFALKVT PFSVKLSSEA SVAVDRFKSL FLMEVHMIK 480
 NKPASPHIKI EMKNATLAWD SSHSSIQNSP KLTPKMKKDX RASRGKKEKV RQLQTEHQA 540
 VLAQKQHLH LMSDERPSPE EEEGKHILG HLRLQRTLHS IDLEIQEGLK VGICGSVSGG 600
 TSLISAILG QMTILGEGSIA ISGTFAYVAQ QANILNATLR DNILFGKEYD EERYNSVLNS 660
 CCLRPLDALI PSSDLTEIGE RANLSSGGQR QRISLARALY SDRSIYILD PLSALDAHV 720
 NHIFNSAIRK PLKSKTVLTV THQLQYLVD C DEVIFMKEGC ITERGTHEEL MNLNGDYATI 780
 FNNLLGETP HVEINSKKTET SGSQKKSQDK GPKTGSVKKE KAVKPEBQGL VQLEBKGGGS 840

WO 02/086443

PCT/US02/12476

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

VFHSVYGVYI QAAGGPIAFL VIMALFMLNV GSTAFSTWML SYWIKQSGN TTVTRGNETS 900
VSDSMKDNPH MQYYASIAL SMAVMLILKA IRGVVFKGT LRASSRLHDE LFRILRSPM 960
KFDDTTPTGR ILNRFSDMD EVDVRLPFQA EMFIONVILV PFCVGMIAV FFWFLVAVGP 1020
LVILFSLVHI VSRVLIRELK RLDNITQSPF LSHITSSIQG LATIHAYNKG QEFHLRYQEL 1080
LDDNQAFFL PTCAMRNLA VRLDLISIALI TTTGLMIVLM HQIIPAYAG LAISYAVQLT 1140
GLFOFTVRLA SETEARTSV ERINHYIKTL SLEAPARIKN KAPSPDWPE GEVTFENAEM 1200
RYRENLPVLV KKVSPFIKPK EKIGIVGRGT SGKSSSLGML FRLVELSGGC IKIDGVRISD 1260
IGLADLRSLK SIIPQEPVLF SGTVRSNLD PNOYTEDQIN DALERTHEKE CIAQLPLKLE 1320
SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETDLIQE TIREAFADCT 1380
MLTIAHRLHT VLGSDRIMVL AQGQVVEFDT PSVLLSNDSS RFYAMFAAAE NKVAVKG

Seq ID NO: 586 DNA sequence
Nucleic Acid Accession #: NM_001327.1
Coding sequence: 89..631

1 11 21 31 41 51
AGCAGGGGGC GCTGTGTGTA CCGAGAATAC GAGAATACCT CGTGGGCCCT GACCTTCTCT 60
CTGAGAGCCG GGCAGAGGCT CCGAGCCAT GCAGGCCGAA GGCCGGGCA CAGGGGGTTC 120
GACGGGCGAT GCTGATGGCC CAGGAGGCC TGGCATTCT GATGGCCAG GGGGCAATGC 180
TGGCGGCCCA GGAGAGCGCG GTGCCACGG CGGCAGAGGT CCCCGGGCG CAGGGGCAGC 240
AAGGGCCTCG GGGCCGGGAG GAGGCGCCCC GCGGGGTCCG CATGCGCGCG CGGCTTCAGG 300
GCTGAATGGA TGCTGCAGAT GCGGGGCCAG GGGCGCCGAG AGCCGCTTGC TTGAGTTCTA 360
CCTGCGCATG CCTTTCGCGA CACCCATGGA AGCAGAGCTG GCCCGCAGGA GCCTGGCCCA 420
GGATGCCCCA CCGCTTCCCG TGCCAGGGGT GCTTCTGAAG GAGTTCACTG TGTCCGGCAA 480
CATACTGACT ATCCGACTGA CTGCTGCAGA CCACCGCCAA CTGCAGCTC CCATCAGCTC 540
CTGTCTCCAG CAGCTTCCCG TGTGATGTG GATCAGCGAG TGCTTCTGCG CCGTGTITTT 600
GGCTCAGCTC CCTCAGGGC AGAGGCGCTA AGCCGACCT GCGCGCCCTT CTTAGGTCAT 660
GCCTCCTCCC CTAGGGAATG GTCCAGCAC GAGTGGCCAG TTCATTGTG GGGCCTGATT 720
GTTTGTGCTG GGAGGAGGAC GGCTTACATG TTGTGTTCTG TAGAAAATAA AACTGAGCTA

Seq ID NO: 587 Protein sequence
Protein Accession #: NP_001318.1

1 11 21 31 41 51
MQAEGRGTTG STGDADGPGG PGIPDGPGGN AGGPGEAGAT GGRGPRGAGA ARASGPGGGA 60
PRGPHGGAAS GLNGCCRCGA RGPESRLLEF YLAMPFATPM EALARRSLA QDAPFLPVFG 120
VLLKEFTVSG NLTIRLTAA DHRQLQLSIS SCLQLSLLM WITQCLFVFP LAQPPSGQRR

Seq ID NO: 588 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 52..459

1 11 21 31 41 51
CCTCGTGGGC CCTGACCTTC TCTCTGAGAG CCGGGCAGAG GCTCCGGAGC CATGCAGGCC 60
GAAGGCCAGG GCACAGGGGG TCGACGGGC GATGCTGATG GCCAGGAGG CCTTGGCATT 120
CCTGATGGCC CAGGGGGCAA TGCTGGCGGC CCAGGAGAGG CGGGTCCAC GGGCGGCAGA 180
GGTCCCGCGG GCGCAGGGGC AGCAAGGGCC TCGGGGCCGA GAGGAGCGCG CCCCGGGGT 240
CCGATGCGCG GTGCGGCTTC TGCGCAGGAT GGAAGGTGCC CCTGCGGGGC CAGGAGGCGG 300
GACAGCGGCC TGCTTCAGTT CCGACTGACT GCTGCAGACC ACCGCAACT CCACTCTCCC 360
ATCAGCTCCT GTCTCCAGCA GCTTCCCTG TTGATGTGGA TCACGCACTG CTCTCTGCCC 420
GTGTTTGTGG CTCAGGCTCC CTCAGGGCAG AGGCGCTAAG CCAGCCTGG CGCCCTCTCC 480
TAGGTCATGC CTCTCCCTCT AGGGAATGGT CCCAGCACGA GTGGCCAGTT CATTTGTGGG 540
GCCTGATTGT TTGTCGCTGG AGGAGGACGG CTTACATGTT TGTTCCTGTA GAAATAAAG 600
CTGAGCTA

Seq ID NO: 589 Protein sequence
Protein Accession #: Eos sequence

1 11 21 31 41 51
MQAEGRGTTG STGDADGPGG PGIPDGPGGN AGGPGEAGAT GGRGPRGAGA ARASGPRGGA 60
PRGPHGGAAS AQDGRCPGGA RRPDSRLLF RLTAADHRQL QLSISSCLQQ LSLLMWITQC 120
FLPVFLAQAP SGQRR

Seq ID NO: 590 DNA sequence
Nucleic Acid Accession #: NM_005562.1
Coding sequence: 90..3671

1 11 21 31 41 51
ACAGCGGAGC GCAGAGTGAG AACCACCAAC CGAGGCGCGG GGCAGCGACC CCTGCAGCGG 60
AGACAGAGAC TGAGCGGGCC GGCACCGCCA TGCCTGCGCT CTGGCTGGGC TGCTGCTCT 120
GCTTCTCGCT CCTCTGCCCC GCAGCGCGGG CCACCTCCAG GAGGGAAGTC TGTGATTGCA 180
ATGGGAAGTC CAGGCAAGTC ATCTTTGATC GGGAACTTCA CAGACAACT GGTAAATGGAT 240
TCCGTCGCTC CACTGCAAT GACAACACTG ATGGCATTCA CTGCGAGAAG TGCAAGAATG 300
GCTTTTACCG GCACAGAGAA AGGACCGCT GTTTCCTGTT CAATTGTAACT TCCAAAGGTT 360
CTCTTAGTGC TCGATGTGAC AACTCTGGAC GGTGCAGCTG TAAACCAAGT GTGACAGGAG 420
CCAGATGCGA CCGATGTCTC CAGGCTTCC ACATGCTCAC GGATGCGGGG TGCAACCAAG 480
ACAGAGACT GCTAGACTCC AAGTGTGACT GTGACCCAGC TGGCATCGCA GGGCCTGTG 540
ACGCGGGCCG CTGTGTCTGC AAGCCAGCTG TTAAGTGAGA ACGCTGTGAT AGGTGTGCGAT 600
CAGGTTACTA TAATCTGGAT GGGGGGAACC CTGAGGGCTG TACCCAGTGT TTCTGCTATG 660
GGCATTGAGC CAGCTGCCCC AGCTCTGAG AATACAGTGT CCATAAGATC ACCTCTACCT 720
TTATCAAGA TGTGATGGC TGAAGGCTG TCCAACGAAA TGGGTCTCTC GCAAAGCTCC 780
AATGGTCACA GCGCCATCAA GATGTGTTTA GCTCAGCCCA ACGACTAGAC CCTGTCTATT 840

	TTGTGGCTCC	TGCCAAATTT	CTTGGGAATC	AACAGGTGAG	CTATGGGCAA	AGCCTGTCT	900
	TTGACTACCG	TGTGGACAGA	GGAGGCAGAC	ACCCATCTGC	CCATGATGTG	ATTCTGGAAG	960
	GTGCTGGTCT	ACGGATCACA	GCTCCCTTGA	TGCCACTTGG	CAAGACACTG	CCTTGTGGGG	1020
5	TCACCAAGAC	TTACACATTC	AGGTTAAATG	AGCATCCAAG	CAATAATTGG	AGCCCCCAGC	1080
	TGAGTTACTT	TGAGTATCGA	AGGTTACTGC	GGAATCTCAC	AGCCCTCCGG	ATCCGAGCTA	1140
	CATATGAGAA	ATACAGTACT	GGGTACATTG	ACAATGTGAC	CCTGATTTCA	GCCCGCCCTG	1200
	TCTCTGGAGC	CCCAGCACCC	TGGGTTGAAC	AGTGTATATG	TCTGTGTGGG	TACAAGGGGC	1260
	AATTTGCGCA	GGATTGTGCT	TCTGGCTACA	AGAGAGATTG	AGCGAGACTG	GGGCTTTTGT	1320
10	GCACCTGTAT	TCCTTGTAA	TGTCAAGGGG	GAGGGGCTGT	TGATCCAGAC	ACAGGAGATT	1380
	GTATTCAGG	GGATGAGAAT	CCTGACATTG	AGTGTGCTGA	CTGCCCAATT	GTTTCTTACA	1440
	ACGATCCGCA	CGACCCCGCG	AGCTGCAAGC	CATGTCCCTG	TCATAACGGG	TTGAGCTGCT	1500
	CAGTGTGCG	GGAGACGGAG	GAGGTGTGTG	GCAATAACTG	CCCTCCCGGG	GTACCGGTG	1560
	CCGCTGTGTA	GCTCTGTGCT	GATGGCTACT	TTGGGGACCC	CTTGTGTGAA	CATGGCCAG	1620
	TGAGGCGCTG	TGACCCCTGT	CAATGCAACA	ACAATGTGGA	CCCCAGTGCC	TCTGGGAATT	1680
15	GTGACCGGCT	GACAGGAGG	TGTTTGAAGT	GTATCCACAA	CACAGCCGGG	ATCTACTGCG	1740
	ACCAAGTCAA	AGCAGGCTAC	TTCCGGGACC	CATTGGCTCC	CAACCCAGCA	GACAAGTGTC	1800
	GAGCTTGCAA	CTGTAAACCC	ATGGGCTCAG	AGCCTGTAGG	ATGTGCAAGT	GATGGCACT	1860
	GTGTTTGCAA	GCCAGGATT	GSTGGCCCCA	ACTGTGAGCA	TGGAGCATT	AGCTGTCCAG	1920
20	CTTGCTATAA	TCAAGTGAAG	ATTGAGATGG	ATCAGTTTAT	GCAGCAGCTT	CAGAGAATGG	1980
	AGGCCCTGGT	TTCAAGGCT	CAGGCTGGTG	ATGGAGTAGT	ACCTGATACA	GAGCTGGAAG	2040
	GCAGGATGCA	GCAGGCTGAG	CAGGCCCTTC	AGGACATTCT	GAGAGATGCC	CAGATTTTCA	2100
	AAGGTGCTAG	CAGATCCCTT	GGTCTCCAGT	TGGCCAAGGT	GAGGAGCCAA	GAGAACAGCT	2160
	ACCAGAGCCG	CCTGGATGAC	CTCAAGATGA	CTGTGGAAG	AGTTCGGGCT	CTGGGAAGTC	2220
25	AGTACCAGAA	CCGAGTTCGG	GATACTCACA	GGCTCATCAC	TCAGATGCAG	CTGAGCCTGG	2280
	CAGAAAGTGT	AGCTTCTCTG	GGAAACACTA	ACATTCCCTG	CTCAGACCCAC	TACGTGGGGC	2340
	CAAAATGGCT	TAAAGTCTG	GCTCAGGAGG	CCACAAGATT	AGCAGAAAGC	CACGTTGAGT	2400
	CAGCCAGTAA	CATGGAGCAA	CTGACAAGGG	AAACTGAGGA	CTATTCCAAA	CAAGCCTCT	2460
	CATGCTGTGC	CAAGGCCCTG	CATGAAGGAG	TGGGAAGCGG	AAGCGGTAGC	CCGACCGGTG	2520
30	CTGTGTGTGA	AGGCTTGTG	GAAAAATTGG	AGAAAAACAA	GTCCCTGGCC	CAGCAGTTGA	2580
	CAGGAGAGGC	TACCAAGCG	GAAATTGAAG	CAGATAGGTC	TTATCAGCAC	AGCTCTCGCC	2640
	TCCTGGATTG	AGTGTCTCGG	CTTCAGGGAG	TCAGTGATCA	GTCCCTTCAG	GTGGAAGAAG	2700
	CAAGAGGAT	CAAAACAAA	GCGGATTAC	TCTCAACGCT	GGAACCCAGG	CATATGGATG	2760
	AGTTCAAGCG	TACACAAAAG	AATCTGGGAA	ACTGGAAAGA	AGAAGCACAG	CAGCTCTTAC	2820
35	AGAATGAAA	AAGTGGGAGA	GAGAAATCAG	ATCAGCTGCT	TTCCCGTGCC	AATCTTGCTA	2880
	AAAGCAGAGC	GACAGAGCAA	CTGAGTATGG	GCAATGCCAC	TTTTATGAAA	GTTGAGAGCA	2940
	TCCTTAAAA	CCTCAGAGAG	TTTGACCTGC	AGGTGGACAA	CAGAAAAGCA	GAAGCTGAAG	3000
	AAGCCATGAA	GAGACTCTCC	TACATCAGCC	AGAAGGTTTC	AGATGCCAGT	GACAAGACCC	3060
	AGCAAGCAGA	AAGAGCCCTG	GGGAGCGCTG	CTGCTGATGC	ACAGAGGGCA	AAGAATGGGG	3120
40	CCGCGGAGGC	CCTGGAATTC	TCCAGTGAGA	TTGAACAGGA	GATTGGGAGT	CTGAACTTGG	3180
	AAGCCAAATG	GACAGCAGAT	GGAGCCTTGG	CCATGGAAAA	GGGACTGGCC	TCTCTGAAGA	3240
	GTGAGATGAG	GGAGTGGAA	GGAGAGCTGG	AAAGGAAGGA	GCTGGAGTTT	GACACGAATA	3300
	TGGATGCAAT	ACAGATGGTG	ATTACAGAAG	CCCAGAAGGT	TGATACCAGA	GCCAAAGAAC	3360
	CTGGGGTTAC	AATCCAAGAC	ACACTCAACA	CATTAGACGG	CCTCTGTCAT	CTGATGGAAC	3420
45	AGCCTCTCAG	TGTAGATGAA	GAGGGGCTGG	TCTTACTGGA	GCAGAACTT	TCCCGAGCCA	3480
	AGACCCAGAT	CAACAGCCAA	CTGCGGCCCA	TGATGTGAGA	GCTGGAAAGG	AGGGCAGCTC	3540
	AGCAGAGGGG	CCACCTCCAT	TTGCTGGAGA	CAAGCATAGA	TGGGATTCTG	GCTGATGTGA	3600
	AGAATCTGGA	GAACATTAGG	GACAACCTGC	CCCCAGGCTG	CTACAATACC	CAGGCTCTTG	3660
	AGCAACAGTG	AAGCTGCCAT	AAATATTCT	CAACTGAGGT	TCTTGGGATA	CAGATCTCAG	3720
50	GGCTCGGGAG	CCATGTGATG	TGAGTGGGTG	GGATGGGGAC	ATTGGAACAT	GTTTAAATGGG	3780
	TATGCTCAGG	TCAACTGACC	TGACCCCAT	CCTGATCCCA	TGGCCAGGTG	GTTGTCTTAT	3840
	TGCACCATAC	TCCTTGCTTC	CTGATGCTGG	GCAATGAGGC	AGATAGCACT	GGGTGTGAGA	3900
	ATGATCAAGG	ATCIGGACCC	CAAAGAATAG	ACTGGATGGA	AAGACAAACT	GCACAGGCAG	3960
	ATGTTTGCC	CATAATAGTC	GTAAGTGGAG	TCTTGGAAAT	TGGACAAGTG	CTGTTGGGAT	4020
55	ATAGTCAACT	TATCTTTTGA	GTAATGTGAC	TAAAGGAAAA	AACITTGACT	TGCCCCAGGC	4080
	ATGAAATTT	TCTTAATGCT	AGAACAGAGT	GCAACCCAGT	CACACTGTGG	CCAGTAAAAAT	4140
	ACTATTGCGT	CATATTGTCC	TCTGCAAGCT	TCTTGCTGAT	CAGAGTTCTT	CCTACTTACA	4200
	ACCCAGGGTG	TGAACATGTT	CTCCATTTTC	AAGCTGGAAG	AAGTGAGCAG	TGTTGGAGTG	4260
	AGGACCTGTA	AGGCAGGCCC	ATTGAGAGCT	ATGGTGCTTG	CTGGTGCCCTG	CCACCTTCAA	4320
60	GTTCTGGACC	TGGGATGAC	ATCCTTTCTT	TAAATGATGC	CATGGCAACT	TAGAGATTGC	4380
	ATTTTATGAT	AGCATTCTC	TACCAGCAAA	GCAAAATGTT	GGAAAGTATT	TACTTTTTCG	4440
	GTTTCAAGT	GATAGAAAAG	TGTGGCTTGG	GCAITGAAAG	AGGTAATAAT	CTCTAGATTT	4500
	ATTAGTCCTA	ATTCAATCCT	ACTTTTCCAA	CACCAAAAAT	GATGCGCATC	AATGTATTTT	4560
	ATCTTATTTT	CTCAATCTCC	TCTCTCTTTC	CTCCACCCAT	AATAAGAGAA	TGTTCTTACT	4620
65	CACACTTCAG	CTGGGTGACA	TCCATCCCTC	CATTATCCTT	TCCATCCATC	TTTCCATCCA	4680
	TACCTCCAT	CCATCCTTCC	AACATATATT	TATTGAGTAC	CTACTGTGTG	CCAGGGGCTG	4740
	GTGGGACAGT	GGTGACATAG	TCTCTGCCCT	CATAGAGTTG	ATTGTCTAGT	GAGGAAGACA	4800
	AGCATTTTTA	AAAAATAAAT	TAAACTTAC	AAACTTTGTT	TGTCAAGT	GGTGTTTATT	4860
	GCAATAACCG	CTTGGTTTGC	AACCTCTTTG	CTCAACAGAA	CATATGTTGC	AAGACCCCTC	4920
70	CATGGGGGCA	CTTGAGTTT	GGCAAGGCTG	ACAGAGCTCT	GGGTTGTGCA	CATTTCCTTG	4980
	CATTCAGCT	GTCACCTCTG	GCCTTTCTAC	AACAGACTGT	TGAGTTATGA		5040
	TAACACCAGT	GGGAATGTCT	GGAGGAACCA	GAGGCACCTC	CACCTGGCT	GGGAAGACTA	5100
	TGGTGCTGCC	TGCTTCTGCT	ATTTCTCTGG	ATTTCTCTGA	AAGTGTTTT	AAATAAGAA	5160
	CAATTGTTAG	ATGCC					

Seq ID NO: 591 Protein sequence
Protein Accession #: NP_005553.1

1	11	21	31	41	51	
MPALWLGCC	LFSLLLPAAR	ATSRREVCD	NGKSRQCIF	RELHROTNG	FRCLNCDNT	60
DGIHCECKN	GFYRERERD	CLPCNCSKG	SLSARCDNSG	RCSCKPVG	ARCDRLPGF	120
HMLTDAGCT	DQRLDLSKD	CDPAGIAGP	DAGRCVCKPA	VTGERCDRC	SGVYNLDGN	180
PBSCTQCFY	GHSASCRSSA	EYSVEKITST	FHQDVGWKA	VQRNGSPAKL	QWSQRHQDVP	240
SSAQLRDPVY	FVAPAKFLGN	QVSYGQSL	FDYRVDGR	HPSAHDVLE	GAGLRITAPL	300
MPLKLTLP	LKTYTFLRLN	EHPSNWSPO	LSYFEYRLL	RNLTLRIRA	TYGEYSTGYI	360
DNVTLSARP	VSGAPAPWVE	QCICPVGYKG	QFCQDCASGY	KRDSARLGP	GTICPCNCG	420
GGACDPDGD	CYSGDENPDI	ECADCPGIPY	NDPDPDRSCK	PCPCHNGFSC	SVMPTTEEVV	480

5
10
15

```

CNCNPPGVTS ARCELCAADSY FGDPFGEHGP VRPCQPCQCN NNVDPSASGN CDRLTGRCLK 540
CIHBTAGIYC DQCKAGIYGD PLAPNPADKC RACNCPNMGs EPVGCERSDGT CVCKPGFGGP 600
NCEHGAFSCP ACYNOVKIQM DQFMQQLQRM EALISKAQCG DGVVPDTELE GRMQQAEQAL 660
QDILRDAQIS EGASRSGLGQ LAKVRSQENS YQSRLLDDLKM TVERVRLGGS QYQNRVRDTH 720
RLITQMQLSL AESEASLGNT NIPASDHYVG PNGFKSLAQE ATRLAESHVE SASNMEQLTR 780
ETEDYSKQAL SILVRALHEG VSGSGSPDGP AVVQGLVEKL ERTKSLAQOL TREATQAEIE 840
ADRSYQHSIR LLDVSRLQGG VSDQSPQVEE AKRIKQKADS LSTLVTRHMD EPKRTQKNLG 900
NNKKEEAQQLL QNGKSGREKS DQLLSRANLA KSRAQEALSM GNATFYEVES ILKNLREFDL 960
QVDNRKAEAE EAMKRLSYIS QKVSADSKT QQAERALGSA AADAQRAKNG AGEALEISSE 1020
IEQEIIGSLNL EANVTADGAL AMEKGLASLK SEMREVEGEL ERKELEFDTN MDAVQMVITE 1080
AQKVDTRAKN AGVTIQDTLN TLDGLLHMD QPLSVDEEGL VLLBQKLSRA KTQINSQLRP 1140
MMSELEERAR QQRGHLHLE TSIDGILADV KNLENIRDNL PPGCYNTQAL EQQ

```

Seq ID NO: 592 DNA sequence
Nucleic Acid Accession #: AF101051.1
Coding sequence: 221..856

20
25
30
35
40
45
50
55
60
65
70
75

```

1 11 21 31 41 51
| | | | |
GAGCAACCTC AGCTCTTAGT ATCCAGACTC CAGCGCGGCC CCGGGCGCGG ACCCCAACCC 60
CGACCCAGAG CTTCTCCAGC GCGCGCGCAG CGAGCAGGGC TCCCCGCTTC AACTTCCTCC 120
GCGGGGCCCA GCCACCTTCG GGAGTCCGGG TTGCCCACTC GCAAACCTCT CGCCTTCTGC 180
ACCTGCCACC CCTGAGCCAG CGCGGGCGCC CGAGCGAGTC ATGGCCAAAG CGGGGCTGCA 240
GCTGTGGGGC TTCTATCTCG CCTTCTCTGG ATGGATCGGC GCCATCGTCA GCACTGCCCT 300
CGCCCCAGTG AGGATTACTT CCTATGCCGG CGACAACATC GTGACCGCCC AGGCCATGTA 360
CGAGGGGCTG TGGATGTCTT GCGTGTCCGA GAGCACCGGG CAGATCCAGT GCAAGTCTT 420
TGACTCTCTG CTGAATCTGA GCAGCACATT GCAAGCAACC CGTGCCTTGA TGGTGTGTGG 480
CATCCTCTCT GGAGTGATAG CAATCTTTGT GGCACCGGTT GGCATGAAGT GTATGAAGTG 540
CTTGAAGAGC GATGAGGTGC AGAAGATGAG GATGGCTGTC ATTGGGGGTG CGATATTTCT 600
TCTTCAGAGT GTGCTATTTT TAGTTGCCAC AGCATGGTAT GGCAATAGAA TCGTTCAAGA 660
ATTCTATGAC CCTATGACCC CAGTCAATGC CAGGTACGAA TTTGGTCAGG CTCTCTTCC 720
TGGCTGGGCT GCTGCTCTCT TCTGCTTCTT GGGAGGTGCC CTACTTTGCT GTTCTGTGCC 780
CGAAGAAACA ACCTCTTACC CAACACCAAG GCGCTATCCA AAACCTGCAC CTTCAGCGGG 840
GAAAGACTAC GTGTGACACA GAGGCAAAAG GAGAAATCA TGTGAAACA AACCGAAAAAT 900
GGACATTGAG ATACTATCAT TAACATTAGG ACCTTAGAAT TTTGGGTATT GTAATCTGAA 960
GTATGGTATT ACAAACAACA CAACAACAACA AAAACCCAT GTGTAAATAT ACTCAGTGCT 1020
AAACATGGCT TAATCTTATT TTATCTTCTT TCCTCAATAT AGGAGGGAAG ATTTTACCAT 1080
TTGTATTACT GCTTCCCATC GAGTAATCAT ACTCAATGG GGAAGGGGGT GCTCCTTAAA 1140
TATATATAGA TGTATATATA TACATGTTTT TCTATTAATA ATAGACAGTA AAATACTATT 1200
CCTATTATGT TGATATAGC ATACTAATA TATCTATAA ATAGGTAAAT GTATTAATT 1260
CCATATTGAT GAAGATGTTT ATTGGTATAT TTTCTTTTTC GTCCCTATAT ACATATGTAA 1320
CAGTCAATAA TCATTACTCT TTCTTCTTTC GCTTTGGGTG CCTTTGCCAC AAGACCTAGC 1380
CTAATTTTACC AAGGATGAAT TCTTTCAATT CTTCATGCGT GCGCTTTTCA TATACTTATT 1440
TTATTTTTTA CCATATCTTT ATAGCACTTG CATCGTTATT AAGCCCTTAT TGTCTTTGTG 1500
TTTCATTGGT CTCTATCTCC TGAATCTAAC ACATTTCTAA GCTTACATTT TAGTTTCTAA 1560
AGCCAAGAAG AATTTATTAC AAATCAGAAC TTTGGAGGCA AATCTTTCTG CATGACCAAA 1620
GTGATAAATT CCTGTGACCC TTCCCAACACA ATCCCTGTAC TCTGACCCAT AGCACTCTTG 1680
TTTGGCTTGA AAATATTGTT CCAATTGAGT AGCTGCATGC TGTTCGCCCA GGTGTGTGTA 1740
CACAACTTTA TTGATGAAT TTTTAAGCTA CTTATTCTAT GTTTTATATC CCCCATAACT 1800
ACCTTTTTGT TCCCATTCCC TTAATTGTAT TGTTTTCCCA AGTGAATTA TCATGCGTTT 1860
TATATCTTCC TAATAAGGTG TGGTCTGTTT GTCTGAACAA AGTGTAGAC TTTCTGGAGT 1920
GATAATCTGG TGACAAATAT TCTCTCTGTA GCTGTAAAGC AGTCACTTAA TCTTTCTACC 1980
TCTTTTTTCT ATCTGCCAAA TTGAGATAAT GATACTTAAC CAGTTAGAAG AGGTAGTGTG 2040
AATATTAATT AGTTTATATT ACTCTCATTC TTTGAACATG AACTATGCGT ATGTAGTGTC 2100
TTTATTGTCT CAGCTGGCTG AGACACTGAA GAAGTCACTG AACAAAACCT ACACACGTAC 2160
CTTCATGTGA TTCCTGCTCT TCCTCTCTCT ACCAGTCTAT TTCCACTGAA CAAAACCTAC 2220
ACACATACCT TCATGTGTTT CAGTGCCTTC CTCTCTCTAC CAGTCTATT CCACGTGAACA 2280
AAACCTACGC ACATACCTTC ATGTGGCTCA GTGCGTTCTC CTCTCTACCA GTCTATTTC 2340
ATTCTTTCAG CTGTGTCTGA CATGTTGTG CTCTGTTCCT TTTTAACAAC TGCTCTTACT 2400
TTTCCAGTCT GTACAGAAAT CTATTTCACT TGAGCAAGAT GATGTATGGA AAGGGTGTG 2460
GCACGTGTGT CTGGAGACCT GGATTGTAGT CTGTGTGCTA TCAATCACCG TCTGTGTTG 2520
AGCAAGGCAT TTGGTCTGCT TAAGCTTATT GCTTCATCTG TAAGCGGTGG TTTGTAATTC 2580
CTGATCTTCC CACCTCACAG TGATGTTGTG GGGATCCAGT GAGATAGAAT ACATGTAAGT 2640
GTGGTTTGTG AATTTGAAAA GTGCTATACT AAGGGAAAGA ATTGAGGAAT TAACGTGATA 2700
CGTTTGTGTT TTGCTTTTCA AATGTTTGAA AATAAAAAAA TGTTAAGAAA TGGGTTTCTT 2760
GCCCTAACCA GTCTCTCAAG TGATGAGACA GTGAAGTAAA ATTGAGTGCA CTAACGAAT 2820
AAGATTCTGA GGAAGTCTTA TCTCTGCGAG TGAGTATGGC CCAATGCTTT CTGTGGCTAA 2880
ACAGATGTAA TGGGAAGAAA TAAAAGCCTA CGTGTGGTA AATCCAACAG CAAGGGAGAT 2940
TTTTGAATCA TAATAACTCA TAAGGTGCTA TCTGTTCACT GATGCCCTCA GAGCTCTTGC 3000
TGTTAGCTGG CAGCTGACGC TGCTAGGATA GTTAGTTTGG AAATGGTACT TCATAATAAA 3060
CTACACAAGG AAAGTCAGCC ACCGTGCTCT ATGAGGAATT GGACCTAATA AATTTTAGTG 3120
TGCCCTCCAA ACCTGAGAAAT ATATGCTTTT GGAAGTTAAA ATTTAAATGG CTTTTCGCCAC 3180
ATACATAGAT CTTCATGATG TGTGAGTGTA ATTCATGTG GATATCAGTT ACCAAACATT 3240
ACAAAAAAT TTTATGGCCC AAAATGACCA ACGAAATGT TACAATAGAA TTTATCCAAAT 3300
TTTGATCTTT TTATATTCTT CTACCACACC TGGAAACAGA CCAATAGACA TTTTGGGGTT 3360
TTATAATGGG AATTTGTATA AAGCATTACT CTTTTCATAT AAATGTTTTT TTAATTTAAA 3420
AAAAGGAAAA AAAAAA AAA

```

Seq ID NO: 593 Protein sequence
Protein Accession #: AAD16433.1

85

```

1 11 21 31 41 51
| | | | |
MANAGLQLLG FILAFLGWIG AIVSTALPQW RIYSYAGDNI VTAQAMYEGE WMSCVSQSTG 60
QIQCKVFDLS LNLSSLTQAT RALMVVGILL GVIAIFVATV GNMCKMCLDE DEVQKMRMAV 120
IGGAIFLLAG LAILVATANY GNRIVQSFYD PMTPVNARYE FGQALFTGWA AASLCLLGGA 180
LLCCSCPRKT TSYPTPRPYP KPARSSGKDY V

```


Seq ID NO: 594 DNA sequence
Nucleic Acid Accession #: NM_006180.1
Coding sequence: 352..2820

5
10
15
20
25
30
35
40
45
50
55
60
65
70

1	11	21	31	41	51	
CCCCCATTCG	CATCTAACAA	GGAATCTGCG	CCCCAGAGAG	TCCCGGACGC	CGCCGTCGGG	60
TGCCCGGGGC	GCOCGGCCAT	GCAGCGACGG	CCGCCGCGGA	GCTCCGAGCA	GCGGTAGGCG	120
CCCCCTGTAA	AGCGGTTGCG	TATGCCGGGA	CCACTGTGAA	CCCTGCGGCC	TGCCGGAACA	180
CTCTTCGCTC	CGGACCAGCT	CAGCCTCTGA	TAAGCTGGAC	TGGCACGCC	CGCAACAAGC	240
ACCGAGGAGT	TAAGAGAGCC	GCAAGCGCAG	GGAAGGCTC	CCCGCACGGG	TGGGGGAAAG	300
CGGCCGGTGC	AGCGCGGGGA	CAGGCACTCG	GGCTGGCACT	GGCTGCTAGG	GATGTCGTCC	360
TGGATAAGGT	GGCATGGACC	CGCCATGGCG	CGGCTCTGGG	GCTTCTGCTG	GCTGGTTGTG	420
GGCTTCTGGA	GGGCGCTTTT	CGCCTGTCCC	ACGTCCTGCA	AATGCACTGC	CTCTCGGATC	480
TGGTGCAGCG	ACCCTTCTCC	TGGCATCGTG	GCATTTCGGA	GATTGGAGCC	TAACAGTGTA	540
GATCTCGAGA	ACATCAACGA	AATTTTCATC	GCAAAACAGA	AAAGGTTAGA	AATCATCAAC	600
GAAGATGATG	TTGAAGCTTA	TGTGGGAGTG	AGAAATCTGA	CAATTGTGGA	TTCTGGGATTA	660
AAATTGTGGT	CTCATAAAGC	ATTCTGAAA	AACAGCAACC	TGCAGCAGAT	CAATTTTACC	720
CGAAACAAAC	TGACGAGTTT	GTCTAGGAAA	CATTTCOGTC	ACCTTGACTT	GTCTGAACTG	780
ATCTCTGGTG	GCAATCCATT	TACATGCTCC	TGTGACATTA	TGTGGATCAA	GACTCTCCAA	840
GAGGCTAAAT	CCAGTCCAGA	CACCTCAGGAT	TTGTACTGCC	TGAATGAAAG	CAGCAAGAAT	900
ATTCCCTCGG	CAAACTCGCA	GATACCCAAT	TGTGGTTTGC	CATCTGCAAA	TCTGGGCGCA	960
CTTAACCTGA	CTGTGGAGGA	AGGAAAGTCT	ATCACATTAT	CCTGTAGTGT	GGCAGGTGAT	1020
COGGTTCCTA	ATATGTATTG	GGATGTTGGT	AACCTGGTTT	CCAAACATAT	GAATGAAACA	1080
AGCCACACAC	AGGGCTCCTT	AAGGATAACT	AACATTTTAT	CGATGACAG	TGGGAAGCAG	1140
ATCTCTTGTT	TGGCGGAAA	TCTTGTAGGA	GAAGATCAAG	ATTCTGTCAA	CCTCACTGTG	1200
CACTTTTGAC	CAACTATCAC	ATTTCTCGAA	TCTCCAACTC	CAGACCAACA	CTGGTGCATT	1260
CCATTCTACT	TGAAAGGCAC	CCCAAAACCA	GCGCTTCAGT	GGTTCTATAA	CGGGGCAATA	1320
TTGAATGAGT	CCAAATACAT	CTGTACTAAA	ATACATGTTA	CCAATCACAC	GGAGTACCAC	1380
GGCTGCCTCC	AGCTGGATAA	TCCCACTCAC	ATGAACAATG	GGGACTACAC	TCTAATAGCC	1440
AAGAATGAGT	ATGGGAAGGA	TGAGAAACAG	ATTTCTGCTC	ACTTCATGGG	CTGGCTCGGA	1500
ATTGACGATG	GTGCAAAACC	AAATATCTCT	GATGTAATTT	ATGAAGATTA	TGGAATCGCA	1560
CGAATGACA	TGGGGGACAC	CACGAAACAG	AGTAATGAAA	TCCCTTCCAC	AGACGTCACT	1620
GATAAAACCG	GTCCGGAACA	TCTCTCGGTC	TATGCTGTGG	TGGTGTATGC	GTCTGTGGTG	1680
GGATTGTGCC	TTTTGGTAAAT	GCTGTTTCTG	CTTAAGTTGG	CAAGACACTC	CAAGTTTGGC	1740
ATGAAGGCC	CAGCCTCCGT	TATCAGCAAT	GATGATGACT	CTGCCAGCCC	ACTCCATCAC	1800
ATCTCCAATG	GGAGTAACAC	TOCATCTTCT	TGGAAGGTTG	GCCGAGATGC	TGTCAATTAT	1860
GGAATGACCA	GAGTCCCTGT	CATTGAAAT	CCCCAGTACT	TTGGCATCAC	CAACAGTCAG	1920
CTCAAGCCAG	ACACATTGTT	TCAGCACATC	AAGCGACATA	ACATTGTTCT	GAAAAGGGAG	1980
CTAGGCGAAG	GAGGCTTTGG	AAAAGTGTTC	CTAGCTGAAT	GCTATAACCT	CTGTCTGAG	2040
CAGGACAAGA	TCTTGGTGGC	AGTGAAGACC	CTGAAGGATG	CCAGTGACAA	TGCACCGAAG	2100
GACTTCCACC	GTGAGGCCGA	GCTCCTGACC	AACCTCCAGC	ATGAGCACAT	CGTCAAGTTC	2160
TATGGGTCT	GCGTGGAGGG	CGACCCCTTC	ATCATGGTCT	TTGAGTACAT	GAAGCATGGG	2220
GACCTCAACA	AGTTCTCTAG	GGCACACGGC	CCTGATGCOG	TGCTGATGGC	TGAGGCAAC	2280
CCGCCCAAGG	AACTGACGCA	GTGCGAGATG	CTGCATATAG	CCGAGCAGAT	CGCCGCGGGC	2340
ATGGTCTACC	TGGCGTCCCA	GCACCTCGTG	CACCGGATTT	TGGCCACAG	GAAGTCCCTG	2400
GTCCGGGAGG	ACTTGTCTGT	GAATATCGGG	GACTTTGGGA	TGTCGCGGGA	CGTGTACAGC	2460
ACTGACTACT	ACAGGTCTGG	TGGCCACACA	ATGCTGCCCA	TTGCTGGAT	GCCTCCAGAG	2520
AGCATCATGT	ACAGGAAATT	CACGACGGAA	AGCGAGCTCT	GGAGCCTGGG	GGTCTGTGTT	2580
TGGGAGATTT	TCACCTATGG	CAAAACAGCCC	TGGTACCAGC	TGTCAAAACA	TGAGGTGATA	2640
GAGTGTATCA	CTCAGGGCCG	AGTCTTGAGC	CGACCCCGCA	CGTGCCCCCA	GGAGGTGTAT	2700
GAGCTGATGC	TGGGTGTCTG	GCAGCGAGAG	CCCCACATGA	GGAAGAACAT	CAAGGGCATC	2760
CATACCCCTC	TCAGAACTTT	GGCCAAAGCA	TCTCCGGTCT	ACCTGGACAT	TCTAGGCTAG	2820
GGCCCTTTTC	CCCAGACCGA	TCCTTCCCAA	CGTACTCCTC	AGACGGGCTG	AGAGGATGAA	2880
CATCTTTTAA	CTGCGCTGGG	AGGCCACCAA	GCTGCTCTCC	TTCACTCTGA	CAGTATTAA	2940
ATCAAGAGCT	CCGAGAAGCT	CTCGAGGGAA	CGAGTGTGTA	CTTCTTCATC	CATAGACACA	3000
GTATTGACTT	CTTTTGTGCA	TTATCTCTTT	CTCTCTTTCC	ATCTCCCTTG	GTGTGTCCTT	3060
TTTCTTTTCT	TAAATTTTCT	TTTTTCTTCT	TTTTTTCGTC	TTCCCTGCTT	CAGGATTTCT	3120
ACCCITTTCT	TTGAATCAAT	CTGGCTTCTG	CATTACTATT	AACTCTGCAT	AGACAAAGGC	3180
CTTAACAAC	GTAATTTGTT	ATATCAGCAG	ACACTCCAGT	TTGCCCAACA	CAACTAACAA	3240
TGCCITTTTG	TATTCCTGCC	TTTGATGTGG	ATGAAAAAAA	GGGAAAAACA	ATATTTCACT	3300
TAAACTTTGT	CACCTCTGCT	GTACAGATAT	CGAGAGTTTC	TATGGATTCA	CTTCTATTTA	3360
TTTATTATTA	TTACTGTTCT	TATTGTTTTT	GGATGGCTTA	AGCCTGTGTA	TAAAAAGAA	3420
AACTGTGTT	CAATCTGTGA	AGCCTTTATC	TATGGGAGAT	TAAACCCAGA	GAGAAAGAA	3480
ATTTATTATG	AAACCGCAAT	TGGGAGGAAC	AAAGACAACC	ACTGGGATCA	GCTGGTGTCA	3540
GTCCCTACTT	AGGAAATACT	CAGCAACTGT	TAGCTGGGAA	GAATGTATTG	GGCACCTTCC	3600
CCTGAGGACC	TTTCTGAGGA	GTAAAAAGAC	TACTGGCCTC	TGTGCCATGG	ATGATTCTTT	3660
TCCCATCACC	AGAAATGATA	GCGTGCAGTA	GAGAGCAAAG	ATGGCTT		

Seq ID NO: 595 Protein sequence
Protein Accession #: NP_006171.1

75
80
85

1	11	21	31	41	51	
MSSWIRWHGP	AMARLWGFCH	LUVGFWRAAF	ACPTSCKCSA	SRWCSDPSP	GIVAPPRLEP	60
NSVDPENITE	IFIANQKRLI	IINEDDVEAY	VGLRNLTIIV	SGLKFAVAKA	FLKNSNLQHI	120
NFTRNKLTL	SRKHFRLDL	SELILVGNPF	TCSCDIMWIK	TLQEAQSSPD	TQDLYCLNES	180
SKNIPLANLQ	IPNCGLPSSN	LAAPNLVVEE	GKSITLSCSV	AGDPVPMYMW	DVGNLVSKHM	240
NETSHTQGS	RITNISSDDS	GKQISCVAEN	LVGEDQDSVN	LTVHFAPTIT	FLESPTSDDH	300
WCIPFTVKGN	PKPALQWFYN	GAILNESKYI	CTKIHVNTHT	EYHGCLQLDN	PTHMNGDYT	360
LIANKNEYGHD	EKQISAHFMG	WPGIDDGANP	NYPDVIYEDY	GTAANDIGDT	TNRSEIPIST	420
DVTDKIGREH	LSVYAVVVIA	SVVGFCLVM	LFLKLARHS	KFGMKGPASV	ISNDDDSASP	480
LHHISNGSMT	PSSSGGPPDA	VIIGMTKIPV	IENPOYFGIT	NSQLKPDFTV	QRIKRNIVL	540
KRELGEAGFG	KVFLAECYNL	CPEQDKILVA	VKTLKDASDN	ARKDFHREAE	LLTNLQHEHI	600
VKFGYGVCEG	DPLIMVFYEM	KHGDNLNKL	AHGPDALVMA	EGNPTEELTQ	SQMLHIAQOI	660
AGMVGVLASQ	HFVHRDLATR	NCLVGENLLV	KIGDFGMSRD	VYSTDYELRVG	GHTMLPIRWM	720

PPESIMYRK F TTESDVWSLG VVLWEIFTYG KQPHYQLSN EVIECITQGR VLQRPRTCPQ 780
EVYELMLGCV QREPHMRINI KGHITLQNL AKASPVYLDI LG

Seq ID NO: 596 DNA sequence
Nucleic Acid Accession #: AF410899
Coding sequence: 483..2999

1 11 21 31 41 51
10 GGGAGCAGGA GCCTCGCTGG CTGCTTCGCT CGCGCTCTAC GCGCTCAGTC CCCGGCGGTA 60
GCAGGAGCCT GGACCCAGGC GCGGCGGGCG GCGGTGAGGC GCGGAGGCC GGCCTCGAGG 120
TGCATACCGG ACCCCCATTC GCATCTAACA AGGAATCTGC GCCCAGAGA GTCCCGGACG 180
CCGCGGTGCG GTGCGCGGCG GCGCGGGCCA TGCAGGACG GCGCGCGCGG AGCTCCGAGC 240
15 AGCGGTAGCG CCCCCCTGTA AAGCGGTTCT CTATGCGGG ACCACTGTGA ACCCTGCCGC 300
CTGCCGGAAC ACTCTTCGCT CCGGACCAGC TCAGCCTCTG ATAAGCTGGA CTGCGCACGC 360
CCGCAACAAG CACGAGGAGG TTAAGAGAGC CGCAAGCGCA GGGGCTGGC TGGCTGCTAG 420
GTGGGGGAAA GCGGCGGTTG CAGCGCGGG CAGGCACTC GCGGCTCTGG GGCCTCTGCT 480
GGATGTCGTC CTGGATAAGG TGGCATGGAC CCGCATGCG CAGGCTCTGG GGCCTCTGCT 540
20 GGCCTGTTGT GGCCTCTGCG AGGCGCGCTT TCGCTGTGCC CAGTCTCTGC AAATGCAGTG 600
CCTCTCGGAT CTGGTCGAGC GACCTTCTCT CTGCGATCGT GGCATTCTCG AGATTGGAGC 660
CTAACAGTGT AGATCTCGAG AACATCACCG AAATTTTCAT ATGTGGGACT GAGAAATCTG ACAATTGTGG 720
AAATCATCAA CGAAGATGAT GTTGAAGCTT ATGTGGGACT GAGAAATCTG ACAATTGTGG 780
ATTCTGGATT AAAATTGTGT GCTCATAAAG CATTTCTGAA AAACAGCAAC CTGCAGCACA 840
TCAATTTTAC CGAAACAAA CTGACGAGTT TGTCTAGGAA ACATTTCCGT CACTTGTACT 900
25 TGCTGAACT GATCCTGGTG GGCAATCCAT TTACATGCTC CTGTGACATT ATGTGGATCA 960
AGACTCTCCA AGAGGCTAAA TCCAGTCCAG ACCTCAGGA TTTGTACTGC CTGAATGAAA 1020
GCAGCAAGAA TATTCCTCTG GCAAACTGCG AGATACCCAA TTGTGGTTTG CCATCTGCAA 1080
ATCTGCGCGC ACCTAACCTC ACTGTGGAGG AAGGAAAGT TATCACAATA TCCCTGTAGT 1140
TGGCAGGTGA TCCGGTTCTT AATATGTTAT GGGATGTTGG TAACCTGGTT TCCAAACATA 1200
30 TGAATGAAAC AAGCCACACA CAGGCTCTCT TAAGGATAAC TAACATTTC TCCGATGACA 1260
GTGGGAAGCA GATCTCTGTG GTGGCGGAAA ATCTGTAGG AGAAGATCAA GATCTGTCA 1320
ACCTCACTGT GCATTTTGA CCAACTATCA CATTTCTGAA ATCTCCAACC TCAGACCACC 1380
ACTGGTGCAT TCCATTCACT GTGAAGGACA ACCCAAAACC AGCGCTTCAG TGGTTCTATA 1440
ACGGGGCAAT ATTGAATGAG TCCAAATACA TCTGTACTAA AATACATGTT ACCAATCACA 1500
35 CCGAGTACCA CGGCTGCCCT CAGCTGAGTA ATCCCACTCA CATGAACAAT GGGGACTACA 1560
CTCTAATAGC CAAGAATGAG TATGGGAAGG ATGAGAAACA GATTTCTGCT CACTTCAATG 1620
GCTGGCTTGG AATTGACGAT GGTGCAAAAC CAAATTATCC TGATGTAATT TATGAAGATT 1680
ATGGAATGCG AGCGAATGAC ATCGGGGACA CCACGAACAG AAGTAATGAA ATCCCTTCCA 1740
40 CAGACGTCAC TGATAAAACC GGTGCGGAAC ATCTCTCGGT CTATGCTGTG TGGTGTATTG 1800
CGTCTGTGTT GGGATTTTGC CTTTGTGTA TGTGTTTCT GCTTAAGTTG GCAAGACACT 1860
CCAGTTTGG CATGAAAGAT TTCTCATGTT TTGGATTTGG GAAAGTAAAA TCAAGACAAG 1920
GTGTTGGCCC AGCCTCGGTT ATCAGCAATG ATGATGACTC TGCCAGCCCA CTCCTATACA 1980
TCTCCATGAG GAGTAACACT CCACTCTTCT CGGAAGGTGG CCCAGATGCT GTCATTATTG 2040
45 GAATGACCAA GATCCCTGTC ATTGAAATC CCACTACTT TGGCATCACC AACAGTCAGC 2100
TAGGCGAAGG AGCCTTTGGA AAGGTGTTCC TAGCTGAATG CTATAACCTC TGTCTGAGC 2160
AGGACAGAT CTGCTGGGCA GTGAAGACCC TGAAGGATGC CAGTGACAA GCACGCAAGG 2220
ACTTCACCGC TGAGGCGGAG CTCCTGACCA ACCTCCAGCA TGAGCACATC GTCAGTTCT 2280
ATGGCGTCTG CAGGAGGGG GACCCCTCA TCACTGCTCT TGAGTACATG AAGCATGGG 2340
50 ACCTCAACAA GTTCTCAGG GCACACGGCC CTGATGCGGT GCTGATGGCT GAGGGCAACC 2400
CGCCCAACGA ACTGACGAG TCGCAGATGC TGCATATAGC GCTGATGGCT GAGGGCAACC 2460
TGGTCTACCT GCGCTCCAG CACTTCGTGC ACCGCGATT GGGCCACCAG AACTGCTTGG 2520
TCGGGAGAA CTGCTGGTG AAAATCGGG ACITTTGGAT GTCCCGGAG GTGTACAGCA 2580
CTGACTACTA CAGGCTGCTG GGCACACAA TGTGCCCCAT TCGCTGGATG CCTCCAGAGA 2640
55 GCATCATGTA CAGGAATTC ACAGCGGAA GCGAGCTCTG GAGCCTGGG GTCTGTTGT 2700
GGGAGATTTT CACCTATGCG AAAACAGCCCT GGTACAGCT GTGCCCCAG GAGGTGTATG 2760
AGTGTATCAC TCAGGGCGGA GTCCCTGAGC GACCCCGCAT GTGCCCCAG GAGGTGTATG 2820
AGCTGATGCT GGGGTGCTGG CAGCGAGAGC CCCACATGAG GAAGAATC AAGGGCATCC 2880
ATACCTCTCT TCAGAACTGT GCCAAGCAT CTCCGCTCTA CCTGGACATT CTAGGCTAGG 2940
60 GCCCTTTTCC CCAGACCGAT CCTTCCCAAC GTACTCTCTA GACGGGCTGA GAGGATGAAC 3000
ATCTTTTAA TCAGGCTGGA GGCCACCAAG CTGCTCTCTT TCACTCTGAC AGTATTAACA 3060
TCAAGACTC CGAGAAGCTC TCGAGGGAAG CAGTGTGTAC TTCTTCATCC ATAGACACAG 3120
TATGACTCT TTTTGGCAT TATCTCTTC TCTCTTTCCA TCTCCCTGG TTGTTCTTT 3180
TTCTTTTCTT AAATTTCTT TTTCTCTTT TTTTCTGCT TCCCTGCTC ACGATTCTTA 3240
65 TTCTTTTCTT TGAATCAATC TGGCTCTGCT ATTACTATTA ACTCTGATA GACAAAGGCC 3300
TTAACAACG TAATTTGTGA TATCAGCAGA CACTCCAGTT TGCCCAACC AACTAACAA 3360
GCCTTGTGT ATTCTGCTC TTGATGTGA TGAAAAAAG GAAAAACAA TATTTCACT 3420
AACTTTGTC ACTTCTGCT TACAGATATC GAGAGTTTCT ATGGATTAC TTCTATTAT 3480
70 TTATTATTAT TACTGTTCTT ATTGTTTTG GATGGCTTAA GCCTGTGTAT AAAAAAGAA 3540
ACTTGTGTTT AATCTGTGAA GCCTTTATCT ATGGGAGATT AAAACAGAG AGAAGAAGA 3600
TTTATTATGA ACCGCAATAT GGGAGGAACA AAGACAACCA CTGGGATCAG CTGGTGTAG 3660
TCCCTACTTA GGAATACTC AGCAACTGTT AGCTGGGAAG AATGTATTCT GCACCTTCCC 3720
CTGAGGACCT TTCTGAGGAG TAAAGAGACT ACTGGCTCT GTGCCATGGA TGATTTCTTT 3780
CCCATCACA GAAATGATAG CGTGCACTAG AGAGCAAGA TGGCTTCCG GAGACACAAG 3840
75 ATGGCGCATA GTGTGCTCG ACACAGTTTT GTCTCTGAG GTTGTATGA TAGCACTGGT 3900
TTGTTCTCA AGCGCTATC ACAGAACCTT TGTCAACTTC AGTTGAAAAG AGGTGGATT 4020
ATGTCAGAG CTCATTTCG GGTCAAGTGG GAAAGCC

Seq ID NO: 597 Protein sequence
Protein Accession #: AAL67965.1

1 11 21 31 41 51
85 MSSWIRWHP AMARLWGFV LUVGFWRFAF ACPTSCKCSA SRIWCDSPSP GIVAFPRLEP 60
NSVDPENITE IFIANQKRL EINEDDVEAY VGLRNLITVD SLKFEVAHKA FLKSNLQHI 120
NFTRNKLTSL SRKHFRLDL SELILVGNPF TCSCDIMWIK TLQEAQSSPD TODLYCLNES 180
SKNIPLANLQ IPNCGLPAN LAAPNLVVEE GKSITLSCSV AGDPVPNMYW DVGNLVSKHM 240

NETSHTQSSL RITNISDDSG GKQISCAVEN LVGEDQDSVN LTVHFAPTIT FLESPTSDEH 300
 WCIPTTVKGN PKPALQWFFYN GAILNESKYI CTKIRVITNHT EVHGCLQLDN PTHMNGDYT 360
 LIAKNEYGKD EKQISAHFMG WPIIDDGANP NYPDVVIYEDY GTAANDIGDT TNRSNEIPST 420
 DVTDKTGREH LSVYAVVVIA SVVGFCLLVM LFLKLARHS KFGMKDFSWF GFGKVKSRQG 480
 VGPASVISND DDSASVPLHI SNGSNTSPSS EGGPDVAVIG MTKIPVIENP QYFGITNSQL 540
 KPDTFVQHIK RHNVILKREL GEGAFGRVFL AECYNLCPEQ DKILVAVKTL KDASDNARKD 600
 FEREAELLTN LQHEHIVKPY GVCVECDPLI MVFEYMKHGD LNKFLRAHGP DAVLMAEGNP 660
 PTELTQSQML HIAQQAAGM VYLAHQHFVH RDLATRNCLV GENLIVKIGD FGMSRDVYST 720
 DYIRVGGHTM LPIRWMPPE S IMYRKFTTES DVWSLGVVLM EIFTYGRQPM YQLSNNEVIE 780
 CITQGRVLQR PRTCPQEVYE LMLGCWQREP HMRKNKIGIH TLQNLAKAS PVYLDILG

Seq ID NO: 598 DNA sequence

Nucleic Acid Accession #: AB052906

Coding sequence: 74..814

1 11 21 31 41 51
 AAAACCTTGA GGTGATTCAT CTTCAGGCT CTCCTTCCAT CAAGTCTCTC CTCCCTAGCG 60
 CTCCTGGTCC TTAATGGCAG CAGCGGCGGC TACCAAGATC CTTCTGTGCC TCCCGCTTCT 120
 GCTCCTGCTG TCGCGCTGGT CCGGGCTGGG GCGAGCGGAC CCTCACTCTC TTGTCTATGA 180
 CATCACGCTC ATCCCTAAGT TCAGACCTGG ACCACGGTGG TGTGCGGTTC AAGGCCAGGT 240
 GGATGAAAAG ACTTTTCTTC ACTATGACTG TGGCAACAG ACAGTCACAC CTGTCACTCC 300
 CCTGGGGAAG AAATAAATG TCACAACGGC CTGGAAGACA CAGAACCCAG TACTGAGAGA 360
 GGTGGTGGAC ATACTTACAG AGCAACTGGG TGACATTGAG CTGGAGAATT ACACACCCAA 420
 GGAACCCCTC ACCCTGAGG CCAGGATGTC TTGTGAGCAG AAAGCTGAAG GACACAGCAG 480
 TGGATCTTGG CAGTTCAGTT TCGATGGGCA GATCTTCTCT CTCTTTGACT CAGAGAAGAG 540
 AATGTGGACA ACGGTTTATC CTGGAGCCAG AAAGATGAAA GAAAAGTGGG AGAATGACAA 600
 GGTGTGGGCC ATGTCTCTCC ATTACTTCTC AATGGGAGAC TGTATAGGAT GGCTTGAGGA 660
 CTCTTGTAGT GGCATGGACA GCACCTGGA GCAAGTGA GAGCACCAC TCGCATGCTC 720
 CTCAGGCACA ACCCAACTCA GGGCCACAGC CACCACCTCT ATCCTTTGCT GCCTCTCAT 780
 CATCTCCTCC TGCTTCATCC TCCCTGGCAT CTGAGGAGAG TCCTTTAGAG TGACAGGTTA 840
 AAGCTGATAC CAAAAGGCTC CTGTGAGCAC GGTCTTGATC AAAGTGGCCC TCTGTCTGG 900
 CCAGCTGCCC ACGACCTACG GTGTATGTCC AGTGGGCTCC AGCAGATCAT GATGACATCA 960
 TGGACCCAAT AGCTCAITCA CTGCCTTGAT TCCTTTTGCC AACAATTTTA CCAGCAGTTA 1020
 TACCTAACAT ATTATGCAAT TTCTCTTGG TGCTACCTGA TGGAAATCTC GCACCTAAAG 1080
 TTCTGGCTGA CTAACAAGA TATATCATTT TCTTCTTCT CTTTTGTGTT GGAATATCAA 1140
 GTACTTCTTT GAATGATGAT CTCTTCTTGG CAAATGATAT TGTCAGTAAA ATATCACGCT 1200
 TAGACTTCAG ACCTCTGGGG ATTCTTCCG TGCTCTGAAA GAGAAATTTT AAATATTTTA 1260
 ATAAGAAAAA ATTATATTA ATGATTGTTT CTTTATGTTT TTTATGTTT TGTACTGATA 1320
 TTTAATAAAA GAGTCTATT TCCCAAAAAA AAAAAAAAAA AA

Seq ID NO: 599 Protein sequence

Protein Accession #: BAB61048.1

1 11 21 31 41 51
 MAAAAATKIL LCLPLLLLL GWSRAGRADP HSLCYDITVI PKFRPGPRWC AVQGGVDEKT 60
 FLHYDCGNKT VTFVSPGLKK LNVTTAWKAQ NPVLREVVDI LTELRLDIQL ENYTPKEPLT 120
 LQARMSCEQK AEGHSSGSWQ FSDGQIFLL FDSEKRMWTT VHPGARKMKE KWENDKVVM 180
 SFHYFSMDC IGWLEDFLMG MDSTLEPSAG APLAMSSGTT QLRATATTLI LCCLLIILPC 240
 FILPGI

Seq ID NO: 600 DNA sequence

Nucleic Acid Accession #: NM_001898.1

Coding sequence: 57..482

1 11 21 31 41 51
 GGCTCTCACC CTCTCTCCT GCAGCTCCAG CTTTGTGCTC TGCCCTCTGAG GAGACCATGG 60
 CCCAGTATCT GAGTACCCTG CTGCTCCTGC TGGCCACCTC AGCTGTGGCC CTGGCCTGGA 120
 GCCCAAGGA GGAGGATAGG ATAATCCCGG GTGGCATCTA TAACGCAGAC CTCAATGATG 180
 AGTGGGTACA CGCTGCCCTT CACTTCGCCA TCAGCGAGTA TAACAAGGCC ACCAAAGATG 240
 ACTACTACAG ACGTCGCTG CCGGTACTAA GAGCCAGSCA ACAGACCGTT GGGGGGTGA 300
 ATTACTTCTT CGACGTAGAG GTGGGCCGCA CCATATGTAC CAAGTCCAG CCCAAGTGG 360
 ACACCTGTGC TTCCATGAA CAGCCAGAAC TGCAAGAGAA ACAGTTGTGC TCTTGTGAGA 420
 TCTACGAAGT TCCCTGGGAG AACAGAAGGT CCCTGGTGAA ATCCAGGTGT CAAGAATCCT 480
 AGGGATCTGT GCCAGGCCAT TCGCACCAGC CACCACCCAC TCCACCCCTC TGAGTGTCTC 540
 CCACCCCTGG ACTGGTGCC CCCACCTGCG GGGAGGCTCT CCATGTGCTC TGGCCCAAGA 600
 GACAGACAGA GAAGGCTGCA GGAGTCCTTT GTTGCTCAGC AGGGCGCTCT GCCTCCCTC 660
 CTTCCTTCTT GCTTCTAATA GCCCTGGTAC ATGGTACACA CCCCCCACC TCTGTCAATT 720
 AAACAGTAGC ATGCC

Seq ID NO: 601 Protein sequence

Protein Accession #: NP_001889.1

1 11 21 31 41 51
 MAQYLSTLLL LLATLAVALA WSPKEEDRII PGGIYNADLN DEWVQALHF AISEYNKATK 60
 DDYRRLPLRV LRARQTVGG VNYFFDVEVG RTICTKSQPN LDTCAFHEQP ELQKKQLCSF 120
 EIYEVWENR RSLVKSRCQE S

Seq ID NO: 602 DNA sequence

Nucleic Acid Accession #: NM_003976.2

Coding sequence: 299..961

1 11 21 31 41 51

5
10
15
20
25

CTCTGAGCTT CTCTGAGCCT TGTITGCTCA TCTGGAAAA GGGGATTAAA CCATTACCT 60
CATGGAGTTG TGAAGAATA GCTGCAAAAG ACCTAACACA TAGTAAGGTT CCCAGTGACG 120
CTACTTCTGC TGGGTTGAGT CTAGCTGTGT AGGCCCTCTG TTCTCTCACT GGAGAACTG 180
GGGTGGCAGG CCGGTCCCCC ACAAAGATA ACTCATCTCT TAATTGCAA GCTGCCCTCA 240
CAGGAGGGTG GGGGAACAGC TCAACAATGG CTGATGGGGG CTCTCGGTGT TGATAGAGAT 300
GGAACCTGGA CTTGAGGCC TCTCCACGCT GTCCCACTGC CCTTGGCTA GCGGGCAGCC 360
TGCCCTGTGG CCCACCTGG CCGCTCTGGC TCTGCTGAGC AGGCTCCGAG AGGCTCCCT 420
GGGCTCCGGG CCGCGCAGCC CTGCCCCCGG CGAAGGCCCG CCGCTGTCTT TGGCGTCCCC 480
CGCGCCGAC CTGCGGGGGG GAGCACGGC CCGCTGGTGC AGTGAAGAG CCCGCGGGC 540
GCGCGCCAG CCTTCTGGC CCGCGCCCCG GCGCTGTGCA CCCCCATCTG CTCTTCCCCG 600
GGGGGGCGCG GCGCGCGGGG CTGGGGGGCG GGGCAGCGG GCTCGGGCAG GGGGGGCGG 660
GGGCTGCGGC CTGCGCTGCG AGCTGGTGCC GGTGCGCGCG CTCGCGCTGG GCCACCGCTC 720
CGACGAGCTG GTGCGTTTCC GCTTCTGAGC CGGCTCTGCG CGCGCGCGCG GCTCTCCACA 780
CGACCTCAGC CTGGCGAGCC TACTGGGCGC GGGGGCGCTG CGACCGCCCC CGGCTCCCG 840
GCGCGTCAGC CAGCCCTGCT GCGACCCAC GCGCTAGGAA GCGTCTCTCT TCATGGACGT 900
CAACAGCACC TGGAGAACC TGGACCGCT CTCCGCCACC GCTTGGGCT GCTTGGGCTG 960
AGGGCTGCTT CAGGGCTTTT GCAGACTGGA CCTTACCGG TGGCTCTTCC TGCTTGGGAC 1020
CCTCCGCGAG AGTCCCACTA GCGACGGGCC TCAGCCAGGG ACGAAGGCGT CAAAGCTGAG 1080
AGGCGCCAC CCGTGGGTGA TGGATATCAT CCGGAAACAG GTGAAGGGAC AACTGACTAG 1140
CAGCCCCAGA GCCTCAGCC TCGGATCCG AGCCTAAAG ACACAGAGA CCTCAGCTAT 1200
GGAGCCCTTC GGACCCACTT CTCACAGACT CTGGCACTGG CCAGGCTCG AACCTGGGAC 1260
CCCTCTCTG ATGAACACTA CAGTGGCTGA GGCATCAGCC CCGCGCCAGG CCCTGTAGGG 1320
ACAGCATTTG AAGGACACAT ATTGCAGTTG CTTGGTTGAA AGTGCCTGTG CTGGAACCTG 1380
CCTGTACTCA CTATGGGAG CTGGCCCC

Seq ID NO: 603 Protein sequence
Protein Accession #: NP_003967.1

30
35

1 11 21 31 41 51
MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPVPLAS 60
PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPSALP RGGRAARAGG PGSRARAAGA 120
RGCRLSQLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180
RPVSPQCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 604 DNA sequence
Nucleic Acid Accession #: NM_057091.1
Coding sequence: 783..1445

40
45
50
55
60
65
70
75

1 11 21 31 41 51
ACTGGCCGCT GAGAGAAGAA TCGGGTGGAG CAGAGAGCAG CTGCTGCAGG GCAGACAGCC 60
GGACCCCAAA ATCTGCACGT ACCAGCAGTC AGCCGCCCCA CGCAGGGACC GGCTTACCCG 120
TCGCTCCCCG CCTCACTACA CTTTCTCCCG CCGCTCCGCG GGCTCCCGAG CTCTCTACTT 180
CGCGTGTCTA CAAACTCAAC TCCCGGTTTC CGTGCCTCTC CACCGCTCGA GTTCTCTACT 240
CTCCATATCC GAGGGGCCCC TCCAGCATC TACCCCGCTC CCAACCTCGG GGGACCTAGC 300
CAAGCTAGGG GGGACTGGAT CCGACGGGTG GAGCAGCCAG GTGAGCCCCG AAAGTTGGG 360
CGGGGCGAGG GCGCTCCAG CCCCACCCCG GGATCTGGTG ACGCTGGGGC TGGAAATTGA 420
CACCGGACGG CTGCGGCGCG GGGCAGGAGG CTGCTGAGGG ATGGAGTTGG GCCCGGCCCC 480
CAGACAAGGC CCGGGGGCTC CGCCAGCAGC AGGTCCCTCG GGGCCCGAGC CTCGCTGCCA 540
CCCGGGGCTG GAGCCCCACA CCGGAGGTG CAGACTGGCT GCCAAGGCCA CACTTTTGGC 600
TAAAAGAGGC ACTGCCAGGT GTACAGTCTT GGGCATGGC TGTITGAGCT TCGGGGGAGA 660
GCCCGAGACT GGTCCCGGGA AAGTGTGCTA GAAGAACAAG GTGCAGGACC CCGTGTCTGC 720
TCAACAGAGG GTGGGGGGA CAGCTCAACA ATGGCTGATG GCGCTCCTG GTGTTGATAG 780
AGATGGAATC TGAAGTTGGA GGCCTCTCCA CGCTGTCCCA CTGCCCCGCG CCTAGGCGGC 840
AGCCTGCCCT GTGGCCCAAC CTGGCCGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCT 900
CCCTGGGCTC CGCGCCCCCG AGCCCTGCCC CCGCGGAAG CCGCCCGCT GTCTGGCGT 960
CCCCCGCGCG CCACCTGCCG GGGGGAGCGA CGGCCGCTG GTGCAGTGA AGAGCCCGCG 1020
GGCGCGCGCG GCGAGCTTCT CGGCCGCGCG CCGCGCGCG TGCAACCCCA TCTGCTCTTC 1080
CCCGCGGGGG CCGCGCGGGG CGGCTGGGG GCGCGGGCAG CCGGCTCGG GCAGCGGGGG 1140
CGCGGGGCTG CCGCGTGGC TCGCAGCTGG TCGCGGTGGG CGGCTCGGC CTGGGCCACC 1200
GCTCCGAGCA GCTGGTGGT TTCCGCTTCT GCAGCGGCTC CTGCGCGCG GCGGCTCTC 1260
CACACAGACT CAGCCTGGCC AGCCTACTGG GCGCGGGGC CCTGCGACCG CCGCGGGCT 1320
CCCGGCCGT CAGCCAGCCC TGCTGCGGAC CCACGCGCTA CGAAGCGGTC TCCTTCATGG 1380
ACGTCAACAG CACTGGAGA ACCGTGGACC GCCTCTCCGC CACCGCTGC GGCTGCTGG 1440
GCTGAGGGCT CGCTCCAGGG CTTTGAGAC TGGACCTTA CCGTGGCTC TTCTGCTGCT 1500
GGACCTCCC GCAGAGTCCC ACTAGCCAGC GGCCTCAGCC AGGACGAAG GCCTCAAAGC 1560
TGAGAGGCC CTACCGGTGG GTGATGGATA TCATCCCGCA ACAGGTGAAG GGCAACTGA 1620
CTAGCAGCCC CAGAGCCCTC ACCCTGCGGA TCCAGCCTA AAAGACACCA GAGACCTCAG 1680
CTATGAGGCC CTTGGAACCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTCGAACCTG 1740
GGACCCCTCC TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCGCGCC CAGGCCCTGT 1800
AGGGACAGCA TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAGTGCC TGTGCTGGA 1860
CTGGCCTGTA CTCACTCATG GGAGCTGGCC CC

Seq ID NO: 605 Protein sequence
Protein Accession #: NP_003967.1

80
85

1 11 21 31 41 51
MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPVPLAS 60
PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPSALP RGGRAARAGG PGSRARAAGA 120
RGCRLSQLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180
RPVSPQCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 606 DNA sequence
Nucleic Acid Accession #: NM_057160.1

Coding sequence: 1..714

1 11 21 31 41 51
5 ATGCCCGGCC TGATCTCAGC CGAGGACAG CCCCTCCTTG AGGTCTCTCC TCCCCAAGCC 60
CACCTGGGTG CCTCTTTTCT CCTTGAGGCT CCACTTGGTC TCTCCGCGCA GCTTGCCTTG 120
TGGCCCAACC TGGCCGCTCT GGCTCTGCTG AGCAGCGTCG CAGAGGCTCT OCTGGGCTCC 180
GGCCCCCGCA GCGCTGCCCC CGCGGAAGGC CCCCGGCTTG TCCTGGCTCT CCCCGCGGC 240
CACCTGCGCG GGGGACGACG GCGCGCTGG TGCAGTGGAA GAGCCCGGCG GCGCGCGCG 300
10 CAGCTTCTCT GCGCGCGGCC CCGCGCGCTT GCAACCCCAT CTGCTCTTCC CCGCGGGGGC 360
CGCGCGCGCG GGGTGGGGGG CCGGGGAGCG CGCGCTCGGG CAGCGGGGGC GCGGGGCTGC 420
CGCTCTGCTT CGCAGCTGGT GCGGTGCGCG GCGCTCGGCC TGGGCCACCG CTCGACGAG 480
CTGGTGGGTG TCCGCTTCTG CAGCGGCTCC TGCCGCGCGCG CGCGCTCTCC ACACGACCTC 540
AGCCTGGCCA GCTTACTGGG CCGCGGGGCC CTGCGACGCG CCGCGGGCTC CCGCGCGCTC 600
15 AGCCAGCTCT GCTGCGGAGC CACGCGCTAC GAAGCGGTCT CCTTCATGGA CGTCAACAGC 660
ACCTGGAGAA CCGTGGACCG CCTCTCGGCC ACCGCTCTGG GCTGCTGGG CTGAGGGCTC 720
GCTCCAGGGC TTTGAGAGCT GGACCTTAC CGGTGGCTCT TCCTGCTGG GACCTCCCG 780
CAGAGTCCCA CTAGCCAGCG GCCTCAGCCA GGGACGAAG CCTCAAAGCT GAGAGGCCCC 840
TACCGGTGGG TGATGGATAT CATCCCCGAA CAGGTGAAGG GACAACTGAC TAGCAGCCCC 900
20 AGAGCGCTCA GCTGCGGAGT CCGAGCTTAA AAGACACAG AGACCTCAGC TATGGAGCCC 960
TTCCGAGCCA CTCTCAGAG ACTCTGGCAC TGGCCAGGCC TCGAACTTGG GACCCCTCTC 1020
CTGATGAACA CTACAGTGGC TGAGGCATCA GCCCGCGGCC AGGCCCTGTA GGGACAGCAT 1080
TTGAAGGACA CATATTGCAG TTGCTTGGTT GAAAGTGCCT GTGCTGGAAC TGGCTGTATC 1140
TCACTCATGG GAGCTGGCCC C

Seq ID NO: 607 Protein sequence
Protein Accession #: NP_476501.1

1 11 21 31 41 51
30 MPGLISARGQ PLLEVLPPQA HLGALFLPEA PLGLSAQPAL NPTLAALALL SSVAEASLGS 60
APRSPAPREG PPPVLASPAH HLPGGRTARW CSGRARRPPP QPSRPAPPPP APPSALPRGG 120
RAARAGGPGS RARAAGARGC RLRSQLVFVR ALGLGHRSD ELVRFPCSGS CRRARSPHDL 180
35 SLASLLGAGA LRPPPGSRPV SQPCCRPTRY EAVSFMDVNS TWRTVDRLSA TACGCLG

Seq ID NO: 608 DNA sequence
Nucleic Acid Accession #: NM_057090.1
Coding sequence: 29..715

1 11 21 31 41 51
40 CTGATGGGCG CTCTGCTGTG TGATAGAGAT GGAACCTGGA CTGGAGGCC TCTCCAGCT 60
GTCCCACTGC CCTTGGCCTA GCGGCGAGGC TCCACTTGGT CTCTCCGCGC AGCCTGCCCT 120
45 GTGGCCCAAC CTGGCGGCTC TGGCTCTGCT GAGCAGGCTC GCAGAGGCTC CCTTGGGCTC 180
CGCGCCCGCG AGCCTGCCCC CCGCGGAAGG CCCCCGCTT GTCTTGGCTC CCCCCGCGCG 240
CCACTCTGCG GGGGGAGCGA CCGCCGCTG GTGCACTGGA AGAGCCCGGC GCGCGCGGCC 300
GCAGCCTTCT CCGCCCGCGC CCGCGCGGCC TGCACCCCA TCTGCTCTC CCGCGGGGG 360
CGCGCGGCGC CGGCTGGGGG GCGCGGCGAG CGCGCTCGCG GCAGCGGGG CCGCGGGCTG 420
CGCGCTGCGC TCGCAGCTGG TGCGGCTGCG CGCGCTCGCG CTGGGCCACC GCTCCGACGA 480
50 GCTGGTGGGT TCCGCTTCTT GCAGCGGCTC CTGCGCGCGC GCGCGCTCTC CACACGACCT 540
CAGCCTGGCC AGCTTACTGG GCGCGGGGCG CCTGCGACCG CCGCGGGGCT CCGCGCCGCT 600
CAGCCAGCCC TGCTGCGGAG CACCGCGCTA CGAAGCGGTC TCCTTCTATG ACCTCAACAG 660
CACTGGAGGA ACCGTGGACC GCCTCTCGCG CACCGCTGCG GGCTGCTTGG GCTGAGGGCT 720
CGCTCCAGGG CTCTGACAGC TGGACCTTAA CCGGTGGCTC TTCTGCTCTG GGACCTCTCC 780
55 GCAGAGTCCC ACTAGCCAGC GGCCTCAGCC AGGGACGAAG GCCTCAAAGC TGAGAGGCC 840
CTACCGTGGG GTGATGGATA TCATCCCCGA ACAGGTGAAG GGACAACTGA CTAGCAGCCC 900
CAGAGCCCTC ACCCTGCGGA TCCAGCCTA AAAGACACCA GAGACCTCAG CTATGGAGCC 960
CTTCGGACCC ACTTCTCACA GACTCTGGCA CTGSCCAGGC CTGAACTCTG GGACCCCTCC 1020
TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCGCGCC CAGGCCCTGT AGGGACAGCA 1080
60 TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAAGTGCC TGTGCTGGAA CTGCGCTGTA 1140
CTACTCATG GAGCTGGCCC C

Seq ID NO: 609 Protein sequence
Protein Accession #: NP_476431.1

1 11 21 31 41 51
65 MEIGLGLST LSHCPWPRRQ APLGLSAQPA LNPFLAALAL LSSVAEASLG SAPRSPAPRE 60
GPPPVLASPA GHLPGGRTAR WSGRARRPPP PQSPRPAPPP PAPPSALPRG GRAARAGGPG 120
70 SRARAAGARG RLRSQLVFVR RALGLHRSD ELVRFPCSG SCRRARSPHD LSLASLLGAG 180
ALRPPGSRPV VSQPCCRPTR YEAVSFMDVN STWRTVDRLS ATACGCLG

Seq ID NO: 610 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..1746

1 11 21 31 41 51
80 ATGCCACTGA AGCATTATCT CCTTTTGTG GTGGGCTGCC AAGCCTGGGG TGCAGGGTTG 60
GCTTACCATG GCTGCCCTAG CGAGTGTACC TGCTCCAGGG CCTCCAGGT GGAGTGCAAC 120
GGGGCAGCGA TTGTGGCGGT GCCCACCCCT CTGCGCTGGA ACGCCATGAG CCTGCAGATC 180
CTCAACACGC ACATCACTGA ACTCAATGAG TCCCGGTTCC TCAATATCTC AGCCTCATC 240
GCCCTGAGGA TTGAGAAGAA TGAGCTGTGG CGCATCACGC CTGGGGCCTT CCGAAACCTG 300
GGCTCGCTGC GCTATCTCAG CCTCGCCAAC AACAAGCTGC AGGTTCTGCC CATGGGCTC 360
85 TTCCAGGGCG TGGACGCTCT TGAGTCTCTC CTCTGTGCTA GTAAACAGCT GTTGACATC 420
CAGCCGCGCC ACTTCTCCCA GTGCAGCAAC CTCAAGGAGC TGCAGTTGCA CGGCAACCA 480
CTGGAATACA TCCCTGACGG AGCCTTCGAC CACCTGGTAG GACTCAGAA GCTCAATCTG 540

	GGCAAGAAATA	GCCTCACCCA	CATCTCACCC	AGGGTCTTCC	AGCACCTGGG	CAATCTCCAG	600
	GTCTCTCCGC	TGTATGAGAA	CAGGCTCAGC	GATATCCCCA	TGGGCACITT	TGATGGGCTT	660
	GTAAACCTGC	AGGAACCTGC	TCTACAGCAG	AACCAGATTG	GACTGCTCTC	COCTGGTCTC	720
5	TTCCACAAACA	ACCACAACCT	CCAGAGACTC	TACCTGTCCA	ACAACCACAT	CTCCAGCTG	780
	CCACCAGACA	TCCTCATGCA	GCTGCCCCAG	CTCAACCGTC	TTACTCTCTT	TGGGAATTCC	840
	CTAAGAGAGC	TCTCTCTGGG	GATCTTGGGG	CCCATGCCCA	ACCTGCGGGA	GCTTTGGGTC	900
	TATGACAACC	ACATCTCTTC	TCTACCCGAC	AATGTCTTCA	GCAACCTCCG	CCAGTTGCAG	960
	GTCCCTGATTC	TTAGCCGCAA	TCAGATCAGC	TTCTCTCTCC	CGGGTGCTTT	CAACGGGCTA	1020
10	ACGGAGCTTC	GGGAGCTGTC	CCTCCACACC	AACGCACTGC	AGGACCTGGA	CGGGAATGTC	1080
	TTCCGCATGT	TGGCCAACCT	GCAGAACATC	TCCTTGCGCA	ACAATCGCCT	CAGACAGCTC	1140
	CCAGGGAATA	TCCTCGCCAA	CGTCAATGGC	CTCATGGCCA	TCCAGCTGCA	GAACAACCAG	1200
	CTGGAGAATC	TGCCCTCCGG	CATCTTCGAT	CACCTGGGGA	AACTGTGTGA	GCTGCGGCTG	1260
	TATGACAATC	CCTGGAGGTG	TGACTCAGAC	ATCCTTCCGC	TCCGCAACTG	GCTCTGCTC	1320
	AACCAGCCTA	GGTTAGGGAC	GGACACTGTA	CCTGTGTGTT	TCAGCCAGC	CAATGTCCGA	1380
15	GGCCAGTCCC	TCATTATCAT	CAATGTCAAC	GTTGCTGTTC	CAAGCGTCCA	TGTCCTCGAG	1440
	GTGCTAGTT	ACCCAGAAAC	ACCATGGTAC	CCAGACACAC	CCAGTTACCC	TGACACCA	1500
	TCCGTCTCTT	CTACCACTGA	GCTAACCCAG	CCTGTGGAAG	ACTACACTGA	TCTGACTACC	1560
	ATTACAGTCA	CTGATGACCG	CAGCGTTTGG	GGCATGACCC	AGGCCAGAG	CGGCTGGCC	1620
	ATTGCGGCCA	TTGTAATTGG	CATTGTGCGC	CTGGGCTGCT	CCCTGGCTGC	CTGCGTGGC	1680
20	TGTTGCTGCT	GCAAGAAAGG	GAGCCAAAGT	GTCTGTATGC	AGATGAAGGC	ACCCATGAG	1740
	TGTTAAAGAG	GAGAGCTGGA	GCAGGGCTGG	GGAATGATGG	GACTGGAGGA	CCTGGGAATT	1800
	TCATCTTTCT	GCCTCCACCC	CTGGTCCAT	GGAGCTTTCC	CGTGATTGCT	CTTTCTGGCC	1860
	CTAGATAAAG	GTGTGCTTAC	CTCTTCTGTA	CTTGCCTGAT	TCTCCGCTAG	AGAAGCAGGT	1920
25	CGTGCCGAC	GCTTCTACAA	TCAGGAAGAT	AGATCCAACT	GGCCATGGCA	AAAGCCCTGG	1980
	GGATTTCGGA	TTCTATACCC	TGGGCTTCTT	TCGAGAGGGC	TCTTCTCCCA	AATCTTCCCC	2040
	ACCTGTCTCT	CAAGAACAGC	CTTCCCTGCG	CCCAGGCCCC	CTCCGGGCTT	CTGTAGACTC	2100
	AGTTAGTCCA	CAGCCTGCTC	ACTTCTGTGG	AATAGTTCTC	CGCTGAGATA	GCCCTCTCTG	2160
	CCTAAGTATT	ATGTAAGTTG	ATTTCCCTTC	TTTTGTTTCT	CTTGTGTTG	CTATGGCTTG	2220
30	ACCCAGCATG	TCCCTCAAAA	TGAAAGTTCT	CCCTTGATT	TTCTGTCTCT	GAAGGCAGGG	2280
	TGAGTTCTCT	CTTCAAGAAA	GACTTCAAA	CATTAACTG	GTTCCTTAAG	AGCCGTCAAT	2340
	CAGCCTGGTT	TTGGGGATGC	TATGAAAGAG	AGAAGGAAAA	TCATGCGGCT	CAGTTCTCTG	2400
	AGACAGAAGA	GCGTCATCA	GTGTCTCACT	TGTGATTTTT	ATCTGAAAAA	GGAAGAAAAA	2460
	CCCCAGCACA	GCAAGCTCAG	CCTTTTAGAG	AAGGATATTT	CCAACTGCA	AACTTTGCTT	2520
35	TGAAAGATT	AGCCCTTAA	GGAATGAAAT	CATGTAGAA	TTTGGACTTC	TAAAAACATT	2580
	AAAATCAGCT	TATTAACTAG	GGATAGAGAA	AGAAATCTGG	TGCTGGGGG	TCCTGTGTT	2640
	CACCCCTAGA	GTCTGTTTAA	AAATTTTAA	TTGAAGCATG	TGAAGTGATC	STGCAGAAAA	2700
	GTGGGAACAT	GATAGTGAT	GGCTTGGTGG	ATTTTCAACA	ACTGAACATA	CCTGTGTAAT	2760
	CAGCATCTAG	ACCCAGACCC	AGAGCATCAC	AAATATCCCC	CATCTGGGC	TTTTCCGAGA	2820
40	GGAGATGGGG	GCTTCTGAAG	ATGGACTTAC	CTGGGACCTG	CCCCCATGGA	GCCAGGACGG	2880
	TCCCCCCACA	GTCAGCTGCT	GCAAGGGCCC	CGTGGCCAGG	GGTGGAGGAG	AATATGTGGG	2940
	TGTGGACAGG	ATGGGAGACT	GTGGCCTGAA	CAGGAGATT	TATTATATCT	GGAGACCCCTG	3000
	AGAGACCCCTG	AGACCTGGGG	CACCATGGCT	GGCCAGGTCA	GAAGCATCTC	GACTGCAGAG	3060
	GTCCGTGCAG	CCACACCTCT	TTCCCTGCCA	GCAAGTTGTC	TGCGGCTCAT	CGGAGGCCCC	3120
45	TCCGCTCGGA	GCCTTCTATG	GACGTGATAT	GCCTGTATCT	GTTTTAAATT	TTTATTCTTC	3180
	ACTTAGGGAGA	AGTGAATCG	CTCAGAGATG	AGATCCTTTA	ATTGAAACAG	AAGTGTAAAG	3240
	GAATCTAGTG	GTCTTCTAAT	GTGGTAAAT	TCTCCATCAA	CATCAGATC	AGCTGGCAGC	3300
	TGAATCTCAG	AATCTCACTT	ACAGCAGGCG	ACACGGGGGT	ACACCGATGG	GTCACTAGTG	3360
	TCTTGGGGGG	TCCCTGGAGC	TCTCTCTGG	TGTGGTCTGG	TTAGGAGTTG	AGTTGTTTGC	3420
50	TCCAGGGTTA	TTCTCTCTCT	CGAGTCACAG	TCACACGAA	ACCTGCCTTC	TCTGGCTTTC	3480
	CTGCTATACA	CATATTCA	TGGCGCTCAA	GAAGTTAGGC	TCATGGCAAC	GTGTGCTTTT	3540
	CTCTGGGAGA	GTGCCCCAGT	TTACAGTGAA	ATGGAGAAAT	TCAGGTCTCC	ACGTCTGCCC	3600
	AGGAAAGAAC	TTCACTGAC	TCCACGGGGA	TCTGGAATC	CACGACCAAT	CCCGATCGGC	3660
	TCTTATTAGC	TCCCGCTCC	ACAAGACACC	TGTGCTTTGG	AAATCCACCA	CCAATCCGGA	3720
55	TGCGCTCTTA	TTAGCTCCCC	GCTCCACAAG	ACACCTGTGA	TCTGGAATC	TACCACCAAT	3780
	CCCGATCGGC	TCTTATTAGC	TCCCGCTCC	ACAAGACACC	TGTGACATCC	TCCAGGGCCA	3840
	CAGGAGCAG	TCTTGACCA	TTTTCCCTTC	CAGTTCTCTG	ACAAAAAGTG	TCCAGAGGGC	3900
	TGTTTGCAAA	CACTAGTGCA	CTTTGTAGCT	TTTCAACCTC	TGTCCAGGG	AATCTAGGAG	3960
	AGATGAGGCC	CGTCAGAGTC	AAGAGATGTC	ATCCCCCAG	GGTCTCCAAG	GCATTTCCAC	4020
60	ACTATTGGTG	GCACTGGAG	GACATGCACC	AAGGCTTGCC	AGAGCCAACA	GGAAGTGAGC	4080
	CCAGAGCATG	GCACTAGAGC	ATCACCCGCT	GATGGTGGCC	TGCTGTGGCT	GGTGCCAACA	4140
	GGGGCATCCC	GGCCCGTACC	CCTCCAGACA	GGAAGCATGG	GTTTGCCCCA	AGACCTGTGG	4200
	GGTGCTCCTG	TGAGTGGCCT	CCAGATGTCT	TGTGTCATAG	GCACAAGTGG	GCCAGGGCTG	4260
	GAGGGAGGTG	GGAACCTCA	TCATCCGGTG	GGCCCTGCCA	ATCTTAAACC	AGAACCCTTA	4320
65	GGTATTCCCTG	GCAGTAGCCA	TGACATTGGA	GCACCTTCTT	CTCCAGCCAG	AGGCTGACCT	4380
	GAGGGCCACT	GTCTCTAGAT	GACACCAACC	AGGAGCACCC	TAGGTGAGGG	GTGAGGGCCC	4440
	CCTTATGTGA	ACCTCTTGCC	TCTTCTTTC	TCCCATCAGA	GTGGTTGGAT	GGAGCCATTG	4500
	GCCTCTTTT	CTTCAGCGGG	CCCTTCAACC	TCTCTGCACC	ATGTTGTCTG	GCTGAGGAGC	4560
	TACTAGAAAA	GCTGAGTGGA	GTCTCTTTC	CAACAGATG	ATGCATTGTC	TCAATTCTCA	4620
70	GGGCTGGAAT	GAGCCGGCTG	GTCCCCCAGA	AAGCTGGAGT	GGGGTACAGA	GTTTCAATTT	4680
	CCTCTCTGTT	TACAGTCTCT	TGACAGTCCC	ACGCCATCT	GGAGTGGGAG	CTGGGAGTTA	4740
	GTGTGGGAGA	AGAAACAACA	AAAGCCAATT	AGAACCCTA	TTTTTAAAAA	GTGCTTACTG	4800
	TGCACAGATA	CTCTTCAAGC	ACTGGAAGTG	GATTTCTCT	CTAGCCCTCA	GCACCCCTGC	4860
	GGTAGGAGTG	CCGCTCTAC	CCACTTGTGA	TGGGGTACAG	AGGCACTTGC	TCTTCTGCAT	4920
75	GGTGTCAAT	AGGCTGGGAG	TTTTATTAT	CTCTTCAAC	TTGTACAG	AGCTCATGGC	4980
	TTGTCTTGGG	CTTGTGTCAT	TAAACCAAG	GAAATGGAAG	CCATTCCCTT	GTTGCTCTCC	5040
	TTAGTCTTGG	TCATCAGAAC	CTCACTTGGT	ACCATATAGA	TCAAAAGCTT	TGTAACCA	5100
	GGAAAAAATA	AATCTTCCA	TCCCTTAAAG	AATAGAAATG	TTTGTCCCTC	TCATGGGAAT	5160
	TGGGCTGTAT	GTATATTGTT	CTTCTCTCT	AGAATTTAGA	GATACAAGAG	TCTTACTTAG	5220
80	AACCTTTTCA	GGACACAATT	TCCACAACCT	TTCAGATGCT	GATGTAGAGC	TATTGGGAAA	5280
	GAACCTCCAA	ACTCAGGAAG	TTTGACAGAG	GCAGACAGCT	AGAGATAACT	CGGGACCCAG	5340
	AGTTGGTCCA	CAGATGTTAG	ATGTATCCTA	GCTTTTAGCC	ATAAACCACT	CAAGATTCA	5400
	GCCCCCAGAT	CCACAGTCA	GAACGTAATC	TGCGTTGTTG	GGAAGCCAGC	AGTGGCCTTG	5460
	GGAAAGGAAG	CATGGCTGTG	GTTCAGAGAG	GGTGGGCTGG	CAAGCCACTT	CCGGGGAATA	5520
85	CTCTCTCCGC	CCCAAGTTTC	TTCTTCTCT	AAGGAGAGAT	TGTTCTCACC	AACCCGCTGC	5580
	CTTCATGCTG	CCTTCAAGGC	TAGATCATGT	TTGCTTGTCT	TAGAGAATTA	CTGCAATCA	5640
	GCCCCAGTGC	TTGGCGATGC	ATTTACAGAT	TTCTAGGCCC	TCAGGTTTTC	GTAGAGTGTG	5700
	AGCCCTGGTG	GGCAGGGTGG	GGGGTCTGCT	CTTCTGCTGG	ATGCTGCTTG	TAATCCATT	5760

GGTGTACAGA ATCAACAATA AATAATATAC ATGTAT

Seq ID NO: 611 Protein sequence
Protein Accession #: BAB84587.1

1	11	21	31	41	51	
MPLKHYLLLL	VGCQAWGAGL	AYHGCPSECT	CSRASQVECT	GARIVAVPTP	LPKNAMSLQI	60
LNTHITELNE	SPFLNISALI	ALRIEKNELS	RITPGAFRNL	GSLRYLSLAN	NKLQVLPIGL	120
FGGLDSLESL	LLSSNQLLQI	QPAHFSQCSN	LKELQLHGNH	LEYIPDGAFF	HLVGLTKLNL	180
GKNSLTHISP	RVFOHLGNLQ	VLRLYENRLT	DIPMGTFDGL	VNLQELALQQ	NQIGLLSPGL	240
FHNENLQRL	YLSNNHISQL	PPSIFMQLPQ	LNRLTLFGNS	LKELSLGIFG	PMPNLRLEWL	300
YDNHISLPD	NVFSNLRQLQ	VLILSRNQIS	FISPGAFNGL	TELRELSLHT	NALQDLDDNV	360
FRMLANLQNI	SLQNNRLRQL	PGNIPANVNG	LMAIQLNQNG	LENLPLGIFD	HLGKLCLELRL	420
YDNPHRCDS	ILPLRNWLLL	NQPRLGTDIV	PVCFSPANVR	GQSLIIINVN	VAVPSVHVPE	480
VPSYPETFWY	PDPSPYPDTT	SVSSTTELTS	PVEDYTDLTT	IQVTDDRSMV	GMTQAQSGLA	540
IAAIVIGIVA	LACSLAACVG	CCCCKKRSQA	VLMQMKAPNE	C		

Seq ID NO: 612 DNA sequence
Nucleic Acid Accession #: XM_098151
Coding sequence: 1..447

1	11	21	31	41	51	
ATGATGCATT	TGCTCAATTC	TCAGGGCTGG	AATGAGCCGG	CTGGTCCCCC	AGAAAGCTGG	60
AGTGGGGTAC	AGAGTTCAGT	TTTCCTCTCT	GTTTACAGCT	CCTTGACAGT	CCCACGCCCA	120
TCTGGAGTGG	GAGCTGGGAG	TCAGTGTGTG	AGAAGAAACA	ACAAAGGCCA	ATTAGAACCA	180
CTATTTTAA	AAAGTGCTTA	CTGTGCACAG	ATACTCTTCA	AGCACTGGAC	GTGGATTCTC	240
TCTCTAGCCC	TCAGCACCCC	TGCGGTAGGA	GTCCCGCCTC	TACCCACTTG	TGATGGGGTA	300
CAGAGGCACT	TGCTCTTCTG	CATGGTGTTC	AATAGGCTGG	GAGTTTATT	TATCTCTTCA	360
AACTTTGTAC	AAGAGCTCAT	GGCTTGTCTT	GGCTTTTCTG	CATTAAACCA	AAGGAAATGG	420
AAGCCATTCC	CCTGTGTGTC	TCCTTAG				

Seq ID NO: 613 Protein sequence
Protein Accession #: XP_098151

1	11	21	31	41	51	
MMHLNSQGW	NEPAGPPESW	SGVQSSVFLS	VYSSLTVP RP	SGVGAGSQCW	RRNNKSQLEP	60
LFLKSAICAQ	ILFKHWITWL	SLALSTPAVG	VPPLPTCDGV	QRHLLFCMVF	NRLGVLFISS	120
NFVQELMACL	GLSSLNQRKW	KFPFCCSP				

Seq ID NO: 614 DNA sequence
Nucleic Acid Accession #: NM_002658.1
Coding sequence: 77..1372

1	11	21	31	41	51	
GTCCCGCAG	CGCCGTCGCG	CCCTCCTGCC	GCAGGCCACC	GAGGCCGCGG	CCGTCTAGCG	60
CCCCGACCTC	GCCACCATGA	GAGCCCTGCT	GGCGGCGCTG	CTTCTCTGCG	TCCTGGTCTG	120
GAGCGACTCC	AAAGCGACGA	ATGAAGTCA	TCAAGTTCCA	TGAACTGTG	ACTGTCTAAA	180
TGGAGGAACA	TGTGTGTCCA	ACAAGTACTT	CTCCAACATT	CACTGGTGCA	ACTGCCCAAA	240
GAAATTGGGA	GGGCAGCACT	GTGAAATAGA	TAAGTCAAAA	ACCTGCTATG	AGGGGAATGG	300
TCACTTTAC	CGAGGAAAGG	CCAGCACTGA	CACCATGGGC	CGGCCCTGCC	TGCCCTGGAA	360
CTCTGCACT	GTCCCTTCAGC	AAACGTACCA	TGCCACACAG	TCTGATGCTC	TTCACTGGG	420
CCTGGGGAAA	CATAATTACT	GCAGGAACCC	AGACAACCGG	AGGCGACCTT	GGTGTATGT	480
GCAGGTGGGC	TTAAAGCCGC	TTGTCCAAGA	GTGCATGGTG	CATGACTGCG	CAGATGGAAA	540
AAAGCCCTCC	TCTCTCCAG	AAGAATTAAA	ATTTCACTGT	GGCCAAAAGA	CTCTGAGGCC	600
CGCTTTAAG	ATTATTGGGG	GAGAATTCAC	CACCATCGAG	AACGAGCCCT	GGTTTGCGGC	660
CATCTACAGG	AGGCACCGGG	GGGGCTCTGT	CACCTACGTG	TGTGGAGGCA	GCCTCATCAG	720
CCCTGTCTGG	GTGATCAGCG	CCACACACTG	CTTCATTGAT	TACCCAAAGA	AGGAGGACTA	780
CATCCTCTAC	CTGGGTCTGT	CAAGGCTTAA	CTCCAACACG	CAAGGGGAGA	TGAAGTTTGA	840
GGTGGAAAAC	CTCATCTTAC	ACAAGGACTA	CAGCGCTGAC	ACGCTTGCTC	ACCACAACGA	900
CATTGCCCTG	CTGAAGATCC	GTTCCAAGGA	GGGCAGGTGT	GCGCAGCCAT	CCCGGACTAT	960
ACAGACCATC	TGCCCTGCCCT	CGATGTATAA	CGATCCCCAG	TTTGGCACA	GCTGTGAGAT	1020
CACTGGCTTT	GGAAAAGAGA	ATTCTACCGA	CTATCTCTAT	CCGGAGCAGC	TGAAAATGAC	1080
TGTTGTGAAG	CTGATTTCCT	ACCGGGAGTG	TCAGCAGCCC	CACTACTACG	GCTCTGAAGT	1140
CACCACCAAA	ATGCTATGTG	CTGCTGACCC	CCAATGGAAA	ACAGATTCTT	GCCAGGGAGA	1200
CTCAGGGGGA	CCCCCTGCTT	GTTCCCTCCA	AGGCGCATG	ACTTTGACTG	GAATTGTGAG	1260
CTGGGGCCGT	GGATGTGCCG	TGAAGGACAA	GCCAGGCGTC	TACACGAGAG	TCTCACACTT	1320
CTTACCCTGG	ATCCGCACTG	ACACCAAGGA	AGAGAATGGC	CTGGCCCTCT	GAGGGTCCCC	1380
AGGGAGGAAA	CGGGCACCA	CCGCTTTCTT	GCTGGTTGTC	ATTTTTCAG	TAGAGTCATC	1440
TCCATCAGCT	GTAAGAAGAG	ACTGGGAAGA	TAGGCTCTGC	ACAGATGGAT	TTCCTGTGG	1500
CACCACCAAG	GTGAACGACA	ATAGCTTTAC	CCTCACGGAT	AGGCCTGGGT	GCTGGCTGCC	1560
CAGACCCTCT	GGCCAGGATG	GAGGGGTGGT	CCTGACTCAA	CATGTTACTG	ACCAGCAACT	1620
TGCTTTTTTC	TGGACTGAAG	CCTGCAGGAG	TAAAAAGGG	CAGGGCATCT	CCTGTGCAATG	1680
GGCTCGAAGG	GAGAGCCAGC	TCCCCTGACC	GGTGGGCATT	TGTGAGGCC	ATGGTTGAGA	1740
AATGAATAAT	TTCCCAATTA	GGAGGTGTAA	GCAGCTGAGG	TCTCTTGAGG	GAGCTTAGCC	1800
AATGTGGGAG	CAGCGGTTTG	GGGAGCAGAG	ACACTAACGA	CTTCAGGGCA	GGGCTCTGAT	1860
ATTCCATGAA	TGTATCAGGA	AATATATATG	TGTGTGTATG	TTTGCACT	TGTTGTGTGG	1920
GCTGTGAGTG	TAAGTGTGAG	TAAGAGCTGG	TGCTGTATTG	TTAAGTCTAA	ATATTCTCTT	1980
AAACTGTGTG	GACTGTGATG	CCACACAGAG	TGGTCTTTCT	GGAGAGGTTA	TAGGTCACTC	2040
CTGGGGGCTC	TGGGTGCCCC	CACGTGACAG	TGCTGGGAA	TGTACTTATT	CTGCAGCATG	2100
ACCTGTGACC	AGCACTGTCT	CAGTTTCACT	TTACATAGA	TGTCCCTTTC	TTGGCCAGTT	2160
ATCCCTTCTT	TTTACCTCTAG	TTTATTTTTC	CTTCACTGGG	TGGGGTGAAG	ACCCTCTCTT	2220
ACACTGAATA	TTTATATTTT	ACTATTTTTC	TTTATATTTT	TGTAATTTTA	AATAAAAGTG	2280
ATCAATAAAA	TGTGATTTTT	CTGA				

Seq ID NO: 615 Protein sequence
Protein Accession #: NP_002649.1

```
5      1      11      21      31      41      51
MRALLARLLL CVLVVSDSKG SNEHQVPSN CDCLNGGTCV SNKYFSNIHW CNCPKFKGGQ 60
HCEIDKSKTC YEGNGHFYRG KASTDTMGRP CLPWNSATVL QQTYHAHRSD ALQLGLGKHN 120
YCRNPDNRRR PWCYVQVGLK PLVQECMVHD CADGKKPSSP PEELKFQCGQ KTLRPRFKII 180
GGEFTTIENQ FWFAAIYRRH RGSVTVVCG GSLISPCWVI SATHCFIDYP KKEDYIVYLG 240
RSRLNSNTQG EMKFVEVENLI LHKDYSADTL AHNHDIALLK IRSKEGRCAQ PSRTIQTICL 300
PSMYNDPQFG TSCEITGFGK ENSTDYLYPE QLKMTVVKLI SHRECCQPHY YGSEVTTKML 360
CAADPQWKTD SCQGDGSGGPL VCSLQGRMTL TGI VSWGRGC ALKDKPGVYT RVSHFLPWIR 420
SHTKEENGLA L
```

Seq ID NO: 616 DNA sequence
Nucleic Acid Accession #: NM_024422.1
Coding sequence: 202..2907

```
20      1      11      21      31      41      51
CGCCAAAGGA AAAGCCCTT GGTAGAGAGG CAGGCGCTTC AGAGAAGCTA AGAAAAAGCAC 60
CTCTCCGCGC GCCCCACCTC CTCGCGCTCG CGCTCCTCCT GAGCAGCGGG CCCAGACTGC 120
GCTCCGCGCG CGGCCCTCGC CCGCGGAGC CCTCTACCC CGGCCCGACG CTCGCCCGCG 180
GACCTGCCCG GAGCCCTCTC CATGGAGGCA GCCCGCCCTT TTGCCAGTGA TGCCCTGCAA 300
CTCTGCCCGC TGCTCCTGCT GACCCCTCGC ATCTTAATAT TTGCCAGTGA TGCCCTGCAA 360
AATGTGACAT TACATGTTCG CTCGAACTA GATGCCGAGA AACTTGTGTG TAGAGTTAAC 420
CTGAAAGAGT GCTTTACAGC TGCAAACTA ATTCAATCAA GTGATCCTGA CTTCGAAATT 480
TTGGAGGATG GTTCAGTCTA TACAACAAAT ACTATTCTAT TGTCTCGGA GAAGAGAAGT 540
TTTACCATAT TACTTTCCAA CACTGAGAAC CAAGAAAAGA AGAAAATTCT AAGCGCGGCC 600
GAGCATCAAA CAAAGGTCCT AAAGAAAAGA CATACTAAAG AAAAGTCTCT AAGCGCGGCC 660
AAGAGATGAT GGGCTCCAAT TCCTTGTTCG ATGCTAGAAA ACTCCTTGGG TCCTTTTCCA 720
CTTTTCTTTC AACAGTTTCA ATCTGACACG GCCCAAACT ATACCATATA CTATTCCATA 780
AGAGGTCCTG GAGTTGACCA AGAACCTCGG AATTATTTTT ATGTGGAGAG AGACACTGGA 840
AACTTGTATT GTACTCGTCC TGTAGATCGT GAGCAGTATG AATCTTTTGA GATAATTGCC 900
TTTGCAACAA CTCAGATGGG GTATACTCCA GAACCTCCAC TGCCCTTAAT AATCAAATA 960
GAGGATGAAA ATGATACTA CCCAATTTTT ACAGAAAGAA CTATATCTTT TACAATTTTT 1020
GAAAATTGCA GAGTGGGCAC TACTGTGGGA CAAGTGTGTG CTACTGACAA AGATGAGCCT 1080
GACACGATGC ACACACGCTT GAAGTACTCC ATCATTGGGG AGGTGCCACC ATCACCACCC 1140
CTATTTTCTA TGATCCCAAC TACAGGCGTG ATCACCACAA CATCATCTCA GCTAGACAGA 1200
GAGTTAAATG ACAAGTACCA GTTGAAAATA AAAGTACAAG ACATGGATGG TCAGTATTTT 1260
GGTCTACAGA CAACCTCAAC TTGTATCATT AACATTGATG ATGTAATAGA CCACCTGCCA 1320
ACATTTACTG GTACTTTCTA TGTGACATCA GTGGAAGAAA ATACAGTTGA TGTGGAAATC 1380
TTACGAGTTA CTGTTGAGGA TAAGGACTTA GTGAATCTG CTAACTGGAG AGCTAATATT 1440
ACCATTTTAA AGGCAATGTA AAATGGCAAT TTTAAATTTG TAACAGATGC CAAACCAAT 1500
GAAGGAGTTC TTTGTGTAGT TAAGCCTTTG AATTATGAAG AAAAGCAACA GATGATCTTG 1560
CAAAATGGTG TAGTTAATGA AGCTCCATTT TCCAGAGAGG CTAGTCCAAG ATCAGCCATG 1620
AGCACAGCAT CAGTTACTGT TAATGTAGAA GATCAGGATG AGGGCCCTGA GTGTAAACCT 1680
CCAATACAGA CTGTTCCGAT GAAAGAAAAT GCAGAAAGTG GAAACAACAG CAATGGATAT 1740
AAAGCATATG ACCCAGAAAC AAGAAGTAGC AGTGGCATAA GGTATAAGAA ATTAAGTATG 1800
CAAACAGGGT GGGTCAACAT TGATGAAAAT ACAGGATCAA TCAAAGTTTT CAGAAGCCTG 1860
GATAGAGAGG CAGAGACCAT CAAAATGGGC ATATATAATA TTACAGTCTT TGATCAGAC 1920
CAAGGAGGGA GAACATGTAT GGGGACACTG GGCATTATAC TTCAAGACGT GAATGATAAC 1980
AGCCCATTTA TACCTAAAAA GACAGTATC ATCTGCAAA CCACCATGTC ATCTCGCGAG 2040
ATTGTTGCGG TTGATCTCTA TGAGCCTATC CATGGCCAC CTTTGACTT TAGTCTGGAG 2100
AGTTTACTCT CAGAAGTACA GAGAATGTGG AGACTGAAAG CAATTAATGA TACAGCAGA 2160
CGTCTTCTCT ATCAGAAATG TCCTCCATTT GGCTCATATG TAGTACCTAT AACAGTGAGA 2220
GATAGACTTG GCATGTCTAG TGTCACTTCA TTGGATGTTA CACTGTGTGA CTGCATTACC 2280
GAAAATGACT GCACACATCG TGTAGATCCA AGGATTGGCG GTGGAGGAGT ACAACTTGA 2340
AAGTGGGCCA TCCTTGCAAT ATTGTTGGGC ATAGCAITGC TCTTTTGCA CTGTTTACG 2400
CTGGTCTGTG GGGCTCTGCG GACGTCTAAA CAACCAAAAG TAATTCCTGA TGATTAGCC 2460
CAGCAGAACCC TAATGTATC AAACACAGAA GCTCCTGGAG ATGACAAAGT GTATTCTGCG 2520
AATGGCTTCA CAACCCAAAC TGTGGCGGCT TCTGCTCAGG GAGTTTGTGG CACCGTGGGA 2580
TCAGGAATCA AAAACGAGG TCAGGAGACC ATCGAAATGG TGAAAGGAGG ACACAGACC 2640
TCGGAATCCT CGCGGGGGGC TGGCCACCAT CACACCCTGG ACTCTGCAG GGGAGGACAC 2700
ACGGAGGTGG ACACTGCAG ATACACTTAC TCGGAGTGGC ACAGTTTAC TCAGCCCGCT 2760
CTTGTGAAA AAGTGTATCT GTGTAATCAA GATGAAAATC ACAAGCATGC CCAAGACTAT 2820
GTCCCTGAGC ATAATATGA AGGAAGAGGA TCGGTGGCTG GGTCTGTAGG TTGTTGCA 2880
GAACGACAG AAGAAGATGG GCTTGAATTT TTGGATAATT TGAGGCCCAA ATTTAGGACA 2940
CTAGCAGAA GATGCATGAA GAGATGAGTG TGTCTAATA AGTCTCTGAA AGCCAGTGGC 3000
TTTATGACTT TTAACAAAAA TTACAAACCA AGAATTTTTT AAAGCAGAA AGCTATTG 3060
TGGGGGTTTT TCTCTCATTA TTTGATGGA ATCTCTTGG TCAATGCAC ATTTACAG 3120
AGACACTATA AACAGTACA CAAATTTTTT AATTTTTACA TATTTTTAAA TTACTTATCT 3180
TCTATCCAAG GAGGTCTACA GAGAAATTAA AGTCTGCCTT ATTTGTTACA TTTGGGTATA 3240
ATGACACAG CCAATTTATA GTGCAATAAA ATGTAATTA TTCAAGTCTT TATTATAGAC 3300
TATTTGAAGC ACAACCTAAT GGAAATTTGT AGAGACCTTG CTTTAACTT ATCTCCAGT 3360
AATTAAGTGT TCATGTGGTG CTGGAAGACT GTTGTTTTCC TGAACATCTA AAGTGTGTAG 3420
ACTGCATTCT TGCTATTATT TTATCTTGT AATGTGACCT TTTCACTGTG CAAAGGGAGA
```

Seq ID NO: 617 Protein sequence
Protein Accession #: NP_077740.1

```
85      1      11      21      31      41      51
MEARPSGSW NGALCRLLL TLAILIFASD AKQNVILHVP SKLDAEKLVG RVNLKECPA 60
ANLIHSSDPD FOILEDGSVY TTNTILLSSE KRSFTILLSN TENQEKKIF VFLEHQTKVL 120
```


KKRHTKEKVL RRAKRRWAPI PCSMLENSLG PFPLFLQQVQ SDTAQNTIY YSIRGPGVDQ 180
 EPRNLFYVER DTGNLYCTRP VDREQYESFE IIAFATTPDG YTPELPLPLI IKIEDENDNY 240
 PIFTEETYTF TIFENCVRGT TVGQVCATDK DEPDTRMTRL KYSIIGQVPP SPTLFSMHPT 300
 TGVITTTSSQ LDRELIDKYQ LKIKVQDMG QYFGLQTTST CIINIDVDND HLPFTTRTSY 360
 VTSVEENTVD VEILRVTVED KDLVNTANWR ANYTILKQNE NGNFKIVTDA KTNBGLCVV 420
 KPLNYEEKQO MILQIGVVNE APFSREASPR SAMSTATVTV NVEDQDEGPE CNPPIQTVRM 480
 KENAEVGTTS NGYKAYDPET RSSSGIRYK LTDPTGHWTI DENTGSIKVF RSLDREAETI 540
 KNGIYNITVL ASDQGGRTCT GTLGIILQDV NDNSPFIFPK TVIICKPTMS SAEIVAVDPD 600
 EPIHGPPDFD SLESSTSEVQ RMWRLKAIND TAARLSYQND PPGSYVVPY TVDRILGMSS 660
 VTSLDVTLCD CITENDCTHR VDPRIIGGGV QLQKWLAI LLGIALLFICI LFTLVCGASG 720
 TSKQPKVIPD DLAQONLIVS NTEAPGDDKV YSANGFTTQT VGASAQGVCG TVGSGIKNGG 780
 QETIEMVKGQ HQTSESCRGA GHEHTLDSR GGHTEVDNCR YTYSEWHSFT QPRLGEKVYL 840
 CNQDENHKA QDYVLTYNYE GRGSVAGSVG CCSEKQEDG LEFLDNLEPK FRTLAEACMK 900
 R

Seq ID NO: 618 DNA sequence
 Nucleic Acid Accession #: NM_004949.1
 Coding sequence: 202..2745

20 1 11 21 31 41 51
 CGCCAAAGGA AAAGCCCTT GGATGAGAG CAGGCGCTTC AGAGAAGCTA AGAAAAGCAC 60
 CTCTCCGCGC GCCCCACCTC CTCCGCTCTC GAGCAGCGGG CCCAGACTGC 120
 GCTCCGCGCG CGCCCTCGC CCCGCGGAGC CCTCTACCC CGGCCGACG CTCGCCCGCG 180
 25 GACCTGCCCG GAGCCTCTC CATGGAGGCA GCCGCGCTC CGGCTCTG GAACGGAGCC 240
 CTCTGCCGCG TGCTCTGCT GACCTCTCGG ATCTTAATAT TGCCAGTGA TGCTGCAAA 300
 AATGTGACAT TACATGTTCC CTCCAACTA GATGCCGAGA AACTTGTTGG TAGAGTTAAC 360
 CTGAAGAGAT GCTTTACAGC TGCAAACTA ATTCATTCAA GTGATCCTGA CTGCAAAAT 420
 TTGGAGGATG GTTCAGTCTA TACAACAAAT ACTATTCTAT TGCTCTCGGA GAAGAGAAGT 480
 30 TTTACCATAT TACTTTCCAA CACTGAGAAC CAAGAAAAGA AGAAAATATT TGCTTTTGG 540
 GAGCATCAAA CAAAGGTCTT AAAGAAAAGA CATACTAAAG AAAAGTTCTT AAGGCGCGCC 600
 AAGAGAAGAT GGGCTCCAAT TCCTTGTTGG ATGCTAGAAA ACTCTTGGG TCTTTTCCA 660
 CTTTTCTCTC AACAGGTTCA ATCTGACACG GCCCAAAACT ATACCATATA CTATTCATA 720
 AGAGGTCTCT GAGTTGACCA AGAAGCTCGG AATTTATTTT ATGTGGAGAG AGACACTGGA 780
 35 AACTTGATAT GTACTCTCA CTTAGATCGT GAGCAGTATG AATCTTTTGA GATAATTGCC 840
 TTTGCAACAA CTCCAGATGG GTATACTCCA GAACCTCCAC TGCCCTTAAT AATCAAAATA 900
 GAGGATGAAA ATGATAACTA CCCAATTTTT ACAGAAGAAA CTTATACTTT TACAATTTT 960
 GAAAATGCA GAGTGGGCAC TACTGTGGGA CAAGTGTGTG CTACTGACAA AGATGAGCCT 1020
 GACACGATG ACACACGCTT GAAGTACTCC ATCATTGGGC AGGTGCCACC ATCACCACC 1080
 40 CTATTTCTTA TACCTCCAAC TACAGGCGTG ATCACCACAA CATCATCTCA GCTAGACAGA 1140
 GAGTTAATTG ACAAGTACCA GTTGAAAATA AAAGTACAAG ACATGGATGG TCAGTATTTT 1200
 GGTCTACAGA CAACCTCAAC TTGTATCATT AACATTGATG ATGTAATGTA CCACTTGCCA 1260
 ACATTACTCT GTACTCTTCA TGTGACATCA GTGGAAGAAA ATACAGTTGA TGTGGAATC 1320
 TTACGAGTGA CTGTTGAGGA TAAGGACTTA GTGAATCTG CTAAGTGGAG AGCTAATAT 1380
 45 ACCATTGTTA AGGCAATGCA AAATGGCAAT TTTAAATTTG TAACAGATGC CAAACCAAT 1440
 GAAGGAGTTC TTTGTGTAGT TAAGCCTTTG AATTATGAAG AAAAGCAACA GATGATCTG 1500
 CAAATTTGGT TAGTTAATGA AGCTCCATTT TCCAGAGAGG CTAGTCCAAG ATCAGCCTG 1560
 AGCACAGCAA CAGTTACTGT TAATGTAGAA GATCAGGATG AGGGCCCTGA GTGTAACCT 1620
 50 CCAATACAGA CTGTTGCGAT GAAAGAAAAT GCAGAAGTGG GAACAACAAG CAATGGATAT 1680
 AAAGCATATG AGGCAGAAC AAGAAGTAGC AGTGGCATAA GGTATAAGAA ATTAAGTAT 1740
 CCAACAGGCT GGGTCCCAT TGAAGAAAT ACAGGATCAA TCAAAGTTT CAGAGCCTG 1800
 GATAGAGAGG CAGAGACCAT CAAAAATGGC ATATATAATA TTACAGTCTT TGCATCAGAC 1860
 CAAGGAGGGA GAACATGTAC GGGGACACTG GGCATTATAC TTCAAGACGT GAATGATAAC 1920
 AGCCCATCTA TACCTAAAAA GACAGTATC ATCTGCAAA CCACCATGTC ATCTGCGGAG 1980
 55 ATGTTGCGG TGTATCTCGA TGAGCCTATC CATGGCCAC CTTTGACTT TAGTCTGGAG 2040
 AGTTCTACTT CAGAAGTACA GAGAATGTGG AGACTGAAAG CAATTAATGA TACAGCAGA 2100
 CGTCTTCTCT ATCAGAATGA TCCTCCATTT GGCTCATATG TAGTACCTAT AACAGTGAGA 2160
 GATAGACTTG GCATGTCTAG TGTCACTTCA TTGGATGTTA CACTGTGTGA CTGCATTACC 2220
 60 GAAAATGACT GCACACATCG TGTAGATCCA AGGATTGGCG GTGGAGGAGT ACAACTTGA 2280
 AAGTGGGCCA TCCCTGCAAT ATTGTTGGGC ATAGCAITGC TCTTTGCTAT CCTGTTTACG 2340
 CTGGTCTGTG GGGCTCTCGG GACGCTCTAA CAACCAAAAG TAATCTCTGA TGATTAGCC 2400
 CAGCAGAAC TAAATGTATC AAACACAGAA GCTCCTGGAG ATGACAAAGT GTATTCTGG 2460
 AATGGCTTCA CAACCAAAAC TGTGGGCGCT TCTGCTCAGG GAGTTTGTGG CACCGTGGGA 2520
 65 TCAGGAATCA AAAACGGAGG TCAGGAGACC ATCGAAATGG TGAAGGAGG ACACAGGACC 2580
 TCGGAATCCT GCCGGGGGCG TGGCCACCAT CACACCTGG ACTCCTGCAG GGGAGGACAC 2640
 ACCGAGGTGG ACAACTGCAG ATACACTTAC TCGAGTGGC ACAGTTTAC TCAGCCCGT 2700
 CTTGGTGAAG AATCCATTAG AGGACACACT CTGATTAATA ATTAACAAT GAAAGAAAGT 2760
 GTATCTGTGT AATCAAGATG AAAATCACAA GCATGCCCAA GACTATGTCC TGACATATA 2820
 70 CATGAAGAGA AGAGGATCGG TGGCTGGGTC TGTAGTTTGT TGCACTGAAC GACAAGAAGA 2880
 AGATGGGCTT GAATTTTGGG ATAATTGGGA GCCCAATTT AGGACACTAG CAGAAGCATG 2940
 CATGAAGAGA TGAGTGTGTT CTAATAAGTC TCTGAAAGCC AGTGGCTTTA TGACTTTTAA 3000
 AAAAAATTAC AAACCAAGAA TTTTAAAG CAGAAGATGC TATTGTGGG GGTTTTCTC 3060
 TCATTATTTG TATGGAATCT CTTTGGTCAA ATGCACATTT ACAGAGAGAC ACTATAACA 3120
 75 AGTACACAAA TTTTCAATTT TTTACATATT TTTAAATAC TTATCTTCTA TCCAAGGAGG 3180
 TCTACAGAGA AATTAAAGTC TGCCTTATT GTTACATTG GGTATAATGA CAACAGCCAA 3240
 TTTATAGTGC AATAAAATGT AATTAATCA AGTCTTATT ATAGACTATT TGAAGCACA 3300
 CCTAATGGAA AATTGTAGAG ACCTTGCTTT AACATTATCT CCAGTTAATT AAGTGTTCAT 3360
 GTGGTGCTTG GAAACTGTG TTTTCTGAA CATCTAAAGT GTGTAGACTG CATCTTGCT 3420
 80 ATTATTTTAT TCTGTAAATG TGACCTTTTC ACTGTGCAAA GGGAGATTTC TAGCCAGGCA 3480
 TTGACTATTA CAATTCATT

Seq ID NO: 619 Protein sequence
 Protein Accession #: NP_004940.1

85 1 11 21 31 41 51
 MEAARPSGSW NGALCRLLLL TLAILIFASD ACIGNVTLHVP SKLDAEKLVG RVNLKECF TA 60

ANLIHSSDPD FOILEDGSVY TTNTILLSSE KRSFTILLSN TENQKQKIF VFLEHQTKVL 120
 KKRHTREKVL RRAKRRWAPI PCSMLENSLG PFPLFLOQVQ SDTAQNTYIY YSIRGPGVDQ 180
 EPRNLFYVER DTGNLYCTRP VDREQYESFE IIAFATTPDG YTPELPLPLI IKIEDENDNY 240
 PIPTEETYTF TIFENCVRGT TVGQVCATDK DEPDTHMTRL KYSIIQGVPP SPTLFMSHPT 300
 TGVITTTSSQ LDRELIDKYQ LKIKVQMDG QYFGLQTTST CIINIDVDND HLPTFTRTSY 360
 VTSVEENTVD VEILRVTVED KDLVNTANWR ANYTILKNE NGNFKIVTDA KTNELVLCVV 420
 KPLNYEEKQQ MILQIGVVNE APFSRSASPR SAMSTATVTV NVEDQDEGPE CNPPIQIVRM 480
 KENAEVGTTS NGYKAYDPET RSSSGIRYK LTDPGTWVTI DENTGSIKVF RSLDREAETI 540
 KNGIYNITVL ASDQGGRTCT GTLGIILODV NDNSPPIPKK TVIICKPTMS SAEIVAVDPD 600
 EPIHGPPFDP SLESSTSEVO RMWRLKAIND TAARLSYQND PPFGSYVPI TVRDLGMS 660
 VTSLDVTLCD CITENDCTHR VDPRIIGGGV QLKGWAILAI LLGIALLFICI LFTLVCGASG 720
 TSKQPKVIPD DLAQQNLIVS NTEAPGDDKV YSANGFTTQT VGASAQVCG TVGSGIKNGG 780
 QETIEMVKGQ HQTSESCRGA GHHTLDSR GGHTEVDNCR YTYSEWHSPT QPRLGEESIR 840
 GHTLIKN

Seq ID NO: 620 DNA sequence
 Nucleic Acid Accession #: NM_032545.1
 Coding sequence: 46..718

1 11 21 31 41 51
 AAACATGATCT TCAATGCACT AAGAGAAGGA GACTCTCAAA CCAAAAATGA CCTGGAGGCA 60
 CCATGTCAGG CTTCTGTCTTA CGGTCACTTT GGCATTACAG ATCATCAATT TGGGAAACAG 120
 CTATCAAGA GAGAAACATA ACGGCGGTAG AGAGGAAGTC ACCAAGGTTG CCACTCAGAA 180
 GCACCGACAG TCACCGCTCA ACTGGACCTC CAGTCATTTC GGAGAGGTGA CTGGAGCGC 240
 CGAGGGCTGG GGGCCGGAGG AGCCGCTCCC CTACTCCCGG GCTTTCGGAG AGGGTGGCTC 300
 CGCGCGGCGG CGCTGCTGCA GGAACGGCGG TACCTGCGTG CTGGGSCAGT TCTGCGTGTG 360
 CCGGCGGCGG TACACCGGCC GCTACTGCGA GCATGACCAG AGGGCGCAGT AATGCGGCGC 420
 CCTGGAGCAC GGAGCCTTGA CCCTCCGCGC CTGCCACCTC TGCAAGGTGA TCTTCGGGCG 480
 CCTGCACTGC CTCGCCCTCC AGACGCTGTA CCGCTGTGAC CCGAAAGACT TCCTGGCCTC 540
 CCACGCTCAC GGGCCGAGCG CCGGGGGCGC GCGCAGCGTG CTACTCTTGC TGCCCTGCGC 600
 ACTCCTGCAC CGCTCTCTGC GCGCGGATGC GCGCGCGCAC CCTCGGTCCC TGGTCCCTTC 660
 CGTCTCCAG GGGGAGCGGC GCGCCTCGG AAGCGCGGGA CTGGGCATC GCCTTTAAAT 720
 TTCTATGTTG TAAATAATAG ATGTGTTTAG TTTACGTAA GCTGAAGCAC TGGGTGAATA 780
 TTTTATTGG GTAATAAATA TTTTCATGAA AGCGCCAAAA AAAAAAAAAA AAAAAAAAAA 840
 AAAAAA

Seq ID NO: 621 Protein sequence
 Protein Accession #: NP_115934.1

1 11 21 31 41 51
 MTWRHHVRL FTVSLALQII NLGNSYQREK HNGGREEVTK VATQKHRSQP LNWTSSEHFG 60
 VTGSAGEWGP EEPLPYSRAP GEGASARPRC CRNGGTCVLG SFCVCPAHFT GRYCEHDQRR 120
 SECCALEHGA WTLRACHLCR CIFGALHCLP LQTPDRCDPK DFLASHAHGP SAGGAPSLLL 180
 LLPCALLHRL LRPDAPAHPR SLVPSVLQRE RRPCGRPLG HRL

Seq ID NO: 622 DNA sequence
 Nucleic Acid Accession #: FGENESH predicted
 Coding sequence: 1..390

1 11 21 31 41 51
 ATGAGGTTC A GTGTCTCAG CATGAGGACC GACTACCCCA GAGTGTGCT GGCTCCTGCT 60
 TATGTGTGAG TCTCTCTCCT CCTCTTGTGT CCAAGGGAAG TCATCGCTCC CGCTGGCTCA 120
 GAACCATGGC TGTGCCAGCC GGCACCCAGG TGTGGAGACA AGATCTACAA CCCCTTGGAG 180
 CAGTGTCTGT ACAATGACGC CATCGTGTCC CTGAGCGAGA CCGGCCAATG TGTGCCCTCC 240
 TGCACCTTCT GGCCTCTGCT TGAGCTCTGC TGTCTTGATT CCTTTGGCCT CACAAAGCAT 300
 TTTGTTGTGA AGCTGAAAGT TCAGGGGTGT AATTCCAGT GCCATCATC TCCCATCTCC 360
 AGTAAATGTG AAGAGGCGG GATATGTTAG

Seq ID NO: 623 Protein sequence
 Protein Accession #: FGENESH predicted

1 11 21 31 41 51
 MRFSVSGMRT DYPSVLAPA YVSVCLLLC PREVIAFAGS EPWLCQAPR CGDKIYNPLE 60
 QCCYNDAIVS LSETRQCGPP CTFWPCFELC CLDSFGLTND FVVKLVQGV NSQCHSSPIS 120
 SKCERGRIC

Seq ID NO: 624 DNA sequence
 Nucleic Acid Accession #: M18728.1
 Coding sequence: 51..1085

1 11 21 31 41 51
 GGAGCTCAAG CTCCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60
 CCTCAGCCCC TCCCTGCAGA TTGCATGTCC CCTGGAAGGA GGTCTGCTC ACAGCCCTCAC 120
 TTCTAACTT CTGGAACCCA CCCACCACTG CCAAGCTCAC TATTGAATCC ACGCCATTCA 180
 ATGTCGCAGA GGGGAAGGAG GTTCTTCTAC TCGCCCAAA CCTGCCCCAG AATCGTATTG 240
 GTTACAGCTG GTACAAAGGC GAAAGAGTGG ATGGCAACAG TCTAATTGTA GGATATGTAA 300
 TAGGAACCTA ACAAGCTACC CCAGGGCCCG CATACAGTGG TCGAGAGACA ATATACCCCA 360
 ATGCATCCCT GCTGATCCAG AACGTCAACC AGAATGACAC AGGATTCTAT ACCCTACAAG 420
 TCATAAAGTC AGATCTTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGAGC 480
 TGCCCAAGCC CTCATCTCC AGCAACAAC CCAACCCCGT GGAGGACAAG GATGCTGTGG 540
 CCTTCACTGT TGAACCTGAG GTTCAGAAC CAACCTACCT GTGGTGGTA AATGGTCAGA 600
 GCCTCCCGGT CAGTCCAGG CTGCAGCTGT CCAATGGCAA CATGACCTC ACTCTACTCA 660

5
10
15
20
25
30

GGGTCAAAAG GAACGATGCA GGATCCTATG AATGTGAAAT ACAGAACCCA GCGAGTGCCA 720
 ACCGAGTGGA CCCAGTCACC CTGAATGTCC TCTATGCCCC AGATGTCCCC ACCATTTCCC 780
 CCTCAAGGC CAATTACCGT CCAGGGGAAA ATCTGAACCT CTCCTGCCAC GCAGCCTCTA 840
 ACCCACTGC ACAGTACTCT TGGTTTATCA ATGGGACGTT CCAGCAATCC ACACAAGAGC 900
 TCTTTATCCC CAACATCAGT GTGAATAATA GCGGATCCTA TATGTGCCAA GCCCATAACT 960
 CAGCCACTGG CCTCAATAGT ACCACAGTCA CGATGTGAGT GCTGSCCAGG GTGGCTCTGA 1020
 TCCTCTCAGC TGTGGCCACC GTCCGCATCA CGATTGGAGT GCTGSCCAGG GTGGCTCTGA 1080
 TATAGCAGCC CTGGTGTATT TTCGATATTT CAGGAAGACT GGCAGATTGG ACCAGACCTT 1140
 GAATTCCTCT AGCTCCTCCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGTCTG 1200
 CTCTGCTCCT GAAGCCCTAT ATGCTGGAGA TGGCAAACTC AATGAAAATT TAAAGGGAAA 1260
 ACCCTCAGGC CTGAGGTGTG TGCCACTCAG AGACTTCACC TAACAGAGA CAGTCAAACT 1320
 GCAAAACATG GTGAGAAATT GACGACTTCA CACTATGGAC AGCTTTTCCC AAGATGTCAA 1380
 AACAGACTC CTCATCATGA TAAGGCTCTT ACCCCCTTTT AATTTGTCTT TGCTTATGCC 1440
 TGCTCTTTT GCTTGGCAGG ATGATGCTGT CATTAGTATT TCACAAGAAG TAGCTTCAGA 1500
 GGGTAACCTA ACAGAGTGTG AGATCTATCT TGTCAATCCC AACGTTTAC ATAAAAAAG 1560
 AGATCTCTTA GTGCACCCAG TGACTGACAT TAGCAGCATC TTTAACACAG CGGTGTGTTT 1620
 AAATGTACAG TGGTCTTTT CAGAGTTGGA CTCTAGACT CACCTGTCTT CACTCCCTGT 1680
 TTTAATTCAG CCCAGCCATG CAATGCCAAA TAATAGAATT GCTCCCTACC AATGGGTATC 1740
 GGAGGAGTCT GTGCAGTTTC TGACACTTGT TGTGAAACAT GGCTAAATAC AATGGGTATC 1800
 GCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAATGG CTACACTCAT 1860
 CTGACTCATT CTTTATTCTA TTTTAGTTGG TTTGTATCTT GCCTAAGGTG CGTAGTCCAA 1920
 CTCTTGGTAT TACCTCTCTA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980
 CTCTAAAGC TTTAAATGTC TGCAATGAGC CAGCCATCAA ATAGTGAATG GCTCTCTTTT 2040
 GGCTGGAATT ACAAAACCTA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAGG 2100
 ATAAAGCCC CAAATGGTGG TAACTGATAA TAGCACTAAT GCTTTAAGAT TTGGTCACAC 2160
 TCTCACTAG GTGAGCCGAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAGTGAAGAT 2220
 GTTAAGGAAG AAGATAGATC CAATTAAAAA AAATTAAGAC CAATTAAAAA AAAAAAAGA 2280
 ACCAGGAGA TTCCAGTCTA CTGAGTTAG CATATACAG AAGTCCCCCT TACTTTAAT 2340
 TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTAAACCAA TGTATTATT TCTGTGGTTC 2400
 TGTTCCTTG TTCCAAATTG ACAAACCCA CTGTCTTGT ATTGTATTGC CCAGGGGAG 2460
 CTATCACTGT ACTGTAGAG TGGTGCTGCT TTAATTCATA AATCACAAAAT AAAAGCCAAT 2520
 TAGCTCTATA ACT

Seq ID NO: 625 Protein sequence
 Protein Accession #: AAA59907.1

35
40
45

1 11 21 31 41 51
 MGPPSAPPCR LHPVKEVLL TASLLTFWNP PTTAKLTIES TPFNVAEGKE VLLLAHNLPO 60
 NRIGYSWYKG ERVDGNLIV GYVIGTOQAT PGPAYSGRET IYPNASLLIQ NVTQNDTGPY 120
 TLQVIKSDLV NEEATGQFHV YPELKPISIS SNNSNPVEDK DAVAFTEPE VQNTYLLWV 180
 NGQSLPVSPP LQLSNHMTL TLLSVKRNDA GSYECEIQNP ASANRSDPVT LNVLYGPDVP 240
 TISPSKANR PGENLNLSC AASNPPAQYS WFINGTFQOS TQELFIPNIT VNNSGYSYMCQ 300
 AENSATGLNR TTVTMITVSG SAPVLSAVAT VGITIGVLAR VALI

Seq ID NO: 626 DNA sequence
 Nucleic Acid Accession #: M18728.1
 Coding sequence: 1355..1657

50
55
60
65
70
75
80
85

1 11 21 31 41 51
 GGAGCTCAAG CTCCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60
 CCTCAGCCCC TCCTCTGAGA TTGCATGTCC CCTGGAAGGA GGTCTGTCTC ACAGCCTCAC 120
 TCTAACTCT CTGGAACCCA CCCACCACTG CCAAGCTCAC TATTGAATCC ACCCCATTCA 180
 ATGTGSCAGA GGGGAAGGAG GTTCTTCTAC TCGCCACAAA CCTGCCCCAG AATCGTATTG 240
 GTTACAGCTG GTACAAAGGC GAAAGAGTGG ATGGCAACAG TCTAATTGTA GGATATGTAA 300
 TAGGAACCTA ACAAGCTACC CCAGGGCCCC CATACAGTGG TCGAGAGACA ATATACCCCA 360
 ATGATCCCTT GCTGATCCAG AACGTCAACC AGAATGACAC AGGATTCTAT ACCCTACRAG 420
 TCATAAAGTC AGATCTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGGAGC 480
 TGCCCAAGCC CTCCTCTTCC AGCAACAAC CCAACCCCGT GGAGGACAAG GATGCTGTGG 540
 CCTTCACCTG TGAACCTGAG GTTCAGAAAC CAACCTACCT GTGGTGGGTA AATGGTCAGA 600
 GCCTCCCGGT CAGTCCCGAG CTGCAGCTGT CCAATGGCAA CATGACCCTC ACTCTACTCA 660
 GCGTCAAAAG GAACGATGCA GGATCCTATG AATGTGAAAT ACAGAACCCA CGAGTGCCA 720
 ACCGAGTGGA CCCAGTCACC CTGAATGTCC TCTATGCCCC AGATGTCCCC ACCATTTCCC 780
 CCTCAAGGC CAATTACCGT CCAGGGGAAA ATCTGAACCT CTCTGCCAC GCAGCCTCTA 840
 ACCCACTGC ACAGTACTCT TGGTTTATCA ATGGGACGTT CCAGCAATCC ACACAAGAGC 900
 TCTTTATCCC CAACATCACT GTGAATAATA GCGGATCCTA TATGTGCCAA GCCCATAACT 960
 CAGCCACTGG CCTCAATAGG ACCACAGTCA CGATGTGAGT GCTGSCCAGG GTGGCTCTGA 1020
 TCCTCTCAGC TGTGGCCACC GTCCGCATCA CGATTGGAGT GCTGSCCAGG GTGGCTCTGA 1080
 TATAGCAGCC CTGGTGTATT TTCGATATTT CAGGAAGACT GGCAGATTGG ACCAGACCTT 1140
 GAATTCCTCT AGCTCCTCCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGTCTG 1200
 CTCTGCTCCT GAAGCCCTAT ATGCTGGAGA TGGCAAACTC AATGAAAATT TAAAGGGAAA 1260
 ACCCTCAGGC CTGAGGTGTG TGCCACTCAG AGACTTCACC TAACAGAGA CAGTCAAACT 1320
 GCAAAACATG GTGAGAAATT GACGACTTCA CACTATGGAC AGCTTTTCCC AAGATGTCAA 1380
 AACAGACTC CTCATCATGA TAAGGCTCTT ACCCCCTTTT AATTTGTCTT TGCTTATGCC 1440
 TGCTCTTTT GCTTGGCAGG ATGATGCTGT CATTAGTATT TCACAAGAAG TAGCTTCAGA 1500
 GGGTAACCTA ACAGAGTGTG AGATCTATCT TGTCAATCCC AACGTTTAC ATAAAAAAG 1560
 AGATCTCTTA GTGCACCCAG TGACTGACAT TAGCAGCATC TTTAACACAG CACTCCCTGT 1620
 AAATGTACAG TGGTCTTTT CAGAGTTGGA CTCTAGACT CACCTGTCTT CACTCCCTGT 1680
 TTTAATTCAG CCCAGCCATG CAATGCCAAA TAATAGAATT GCTCCCTACC AGCTGAACAG 1740
 GGAGGAGTCT GTGCAGTTTC TGACACTTGT TGTGAAACAT GGCTAAATAC AATGGGTATC 1800
 GCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAATGG CTACACTCAT 1860
 CTGACTCATT CTTTATTCTA TTTTAGTTGG TTTGTATCTT GCCTAAGGTG CGTAGTCCAA 1920
 CTCTTGGTAT TACCTCTCTA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980
 CTCTAAAGC TTTAAATGTC TGCAATGAGC CAGCCATCAA ATAGTGAATG GTCTCTCTTT 2040
 GGCTGGAATT ACAAAACCTA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAGG 2100
 ATAAAGCCC CAAATGGTGG TAACTGATAA TAGCACTAAT GCTTTAAGAT TTGGTCACAC 2160

WO 02/086443

PCT/US02/12476

TCTCACCTAG GTGAGCGCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAACTGAAAT 2220
GTTAAGGAAG AAGATAGATC CAATTAAAAA AAATTAAGAC CAATTTAAAA AAAAAAAGA 2280
ACACAGGAGA TTCCAGTCTA CTGAGTTAG CATAATACAG AAGTCCCTCT TACTTTAACT 2340
TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTTAACCAA TGTATTATT TCTGTGGTTC 2400
TGTTTCCTTG TTCCAATTG ACAAAACCCA CTGTTCTTGT ATTGTATTGC CCAGGGGGAG 2460
CTATCACTGT ACTGTAGAG TGGTGCTGCT TTAATTCATA AATCACAAAT AAAAGCCAAT 2520
TAGCTCTATA ACT

Seq ID NO: 627 Protein sequence
Protein Accession #: AAA59908.1

1 11 21 31 41 51
MDSFSQDVKT RLLIMIRLFP PPNLSLLMPA SFAWQDDAVI SISQEVASEG NLTECQIYLV 60
NPNVLHKIRD PLVHPVTDIS SIFNTAVCSN VQWSFSELD

Seq ID NO: 628 DNA sequence
Nucleic Acid Accession #: M18728.1
Coding sequence: 2370..2501

1 11 21 31 41 51
GGAGCTCAAG CTCCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60
CCTCAGCCCC TCCTGCGAGA TTGCATGTCC CCTGGAAGGA GGTCTGTCTC ACAGCCTCAC 120
TTCTAACCTT CTGGAACCCA CCCACCACTG CCAAGCTCAC TATTGAATCC ACGCCATTCA 180
ATGTGCGAGC GGGGAAGGAG GTTCTTCTAC TCGCCACAA CCTGCCCCAG AATCGTATTG 240
GTTACAGCTG GTACAAGGC GAAAGAGTGG ATGGCAACAG TCTAATTGTA GGATATGTAA 300
TAGGAACCTA ACAAGCTACC CCAGGGCCCG CATACAGTGG TCGAGAGACA ATATACCCCA 360
ATGCATCCCT GCTGATCCAG AACGTCAACC AGAATGACAC AGGATTCTAT ACCCTACAAG 420
TCATAAAGTC AGATCTTGTG AATGAAGAAG CAACCCGACA GTTCCATGTA TACCCGGAGC 480
TGCCCAAGCC CTCATCTCC AGCAACAAC CCAACCCCGT GGAGGACAAG GATGCTGTGG 540
CCTTCACCTG TGAACCTGAG GTTCAGAAAC CAACCTACCT GTGGTGGGTA AATGGTCAGA 600
GCCTCCCGGT CAGTCCCGAG CTGCAGCTGT CCAATGGCAA CATGACCCCT ACTCTACTCA 660
GGCTCAAAAG GAACGATGCA GGATCCTATG AATGTGAAT ACAGAACCCA GCGAGTGCCA 720
ACCGCAGTGA CCCAGTCACC CTGAATGTCC TCTATGGCCC AGATGTCCCC ACCATTTCCT 780
CCTCAAGAGC CAATTACGCT CCAGGGGAAA ATCTGAACCT CTCTGTCCAC GCAGCCTCTA 840
ACCCACCTGC ACAGTACTCT TGGTTTATCA ATGGGAGCTT CCAGCAATCC ACACAAGAGC 900
TCTTTATCCC CAATCATCT GTGAATAATA GCGGATCCTA TATGTGCCAA GCCCATAACT 960
CAGCCACTGG CTTCAATAGG ACCACAGTCA CGATGATCAC AGTCTCTGGA AGTGTCTCTG 1020
TCCTCTCAGC TGTGGCCACC GTGGGCATCA CGATTGGAGT GCTGGCCAGG GTGGCTCTGA 1080
TATAGCAGCC CTGGTGTATT TCGATATTT CAGGAAGACT GGCAGATTGG ACCAGACCTT 1140
GAATTCCTCT AGCTCTCTCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGTCTG 1200
CTCTGCTCCT GAAGCCCTAT ATGCTGGAGA TGGACAACTC AATGAAAAAT TAAAGGGAAA 1260
ACCTCAGGCG CTTAGGTGTG TGCCACTCAG AGACTTCACC TAACATAGAGA CAGTCAAACT 1320
GCAAAACCAT GTGAGAAATT GAAGACTTCA CACTATGGAG AGCTTTTCCC AAGATGTCAA 1380
AACAAGACTC CTCAATCATG TAAGGCTCTT ACCCCCTTTT AATTGTCTCT TGCTTATGCC 1440
TGCTCTTCTC GCTTGGCAGG ATGATGCTGT CATTAGTATT TCACAAGAAG TAGCTTCAGA 1500
GGGTAACCTA ACAGAGTGTC AGATCTATCT TGTCAATCCC AACGTTTATC ATAAAAAAG 1560
AGATCCTTTA GTGACCCGAG TGACTGACAT TAGCAGCATC TTTAACACAG CCGTGTGTTC 1620
AAATGTACAG TGTCTCTTTT CAGAGTTGGA CTCTAGACT CACCTGTTCT CACTCCCTGT 1680
TTTAATTCAA CCGAGCCCTA CAATGCCAAA TAATAGAATT GCTCCCTACC AGCTGAACAG 1740
GGAGGAGTCT GTGCACTTTC TGACACTTGT TGTGAACAT GGCTAAATAC AATGGGTATC 1800
GCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAAAATG CTACATCAT 1860
CTGACTCAT: CTTATTCTA TTTAGTTGG TTTGTATCTT GCCTAAGGTG CGTAGTCCAA 1920
CTCTTGGTAT TACCCTCTCA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980
CTCTAAAAGC TTTAAATGTC TGCAATGCAG CAGCCATCAA ATAGTGAATG GTCTCTCTTT 2040
GGCTGGAATT ACAAACTCA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAG 2100
ATAAAAGCCC CAAATGGTGG TAACGTATAA TAGCACTAAT GCTTTAAGAT TTGGTCACAC 2160
TCTACCTTAG GTGAGCGCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAACTGAAAT 2220
GTTAAGGAAG AAGATAGATC CAATTAAAAA AAATTAAGAC CAATTTAAAA AAAAAAAGA 2280
ACACAGGAGA TTCCAGTCTA CTGAGTTAG CATAATACAG AAGTCCCTCT TACTTTAACT 2340
TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTTAACCAA TGTATTATT TCTGTGGTTC 2400
TGTTTCCTTG TTCCAATTG ACAAAACCCA CTGTTCTTGT ATTGTATTGC CCAGGGGGAG 2460
CTATCACTGT ACTGTAGAG TGGTGCTGCT TTAATTCATA AATCACAAAT AAAAGCCAAT 2520
TAGCTCTATA ACT

Seq ID NO: 629 Protein sequence
Protein Accession #: AAA59909.1

1 11 21 31 41 51
MLTNVFISVV LPFCSNLTKP TVLVLYCPGG AITVLVEWCC FNS

Seq ID NO: 630 DNA sequence
Nucleic Acid Accession #: NM_016639.1
Coding sequence: 40..429

1 11 21 31 41 51
GGGGGGGGG CAGACAGCGG CGGGGCGAGG ACGTGCACTA TGGCTCGGGG CTCGCTGCGC 60
CGGTGTGCTG GGCTCCTCGT GCTGGGGCTC TGGCTGGCGT TGCTGCGCTC CGTGGCCGGG 120
GAGCAAGCGC CAGGACCGCG CCCCTGTCTC CGGGCGAGCT CTTGAGCGCG GGACCTGGAC 180
AAGTGCATGG ACTGCGCGTC TTGCAGGGCG CGACCGCACA GCGACTTCTG CTTGGGCTGC 240
GCTGCAGCAC CTCTGCGCCC CTTCGGGCTG CTTTGGGCCA TCCTTGGGGG CGCTCTGAGC 300
CTGACCTTCG TGCTGGGGCT GCTTCTGGGC TTTTGGTCTT GGAGACGATG CCGCAGGAGA 360
GAGAAGTTCA CCACCCCAT AGAGGAGACC GCGGAGAGG GCTGCCACG TGTGGCGCTG 420

5 ATCCAGTGAC AATGTGCCCC CTGCCAGCCG GGGCTCGCCC ACTCATCATT CATTCAATCCA 480
 TTCTAGAGCC AGTCTCTGCC TCCAGAGCCG GCGGGGAGCC AAGCTCCTCC AACCACAAGG 540
 GGGGTGGGGG GCGGTGAATC ACCTCTGAGG CCTGGGCCCA GGGTTCAGGG GAACCTTCCA 600
 AGGTGTCTGG TTGCCCTGCC TCTGGCTCCA GAACAGAAAG GGAGCCTCAC GCTGGCTCAC 660
 ACAAACACGC TGACACTGAC TAAGGAACTG CAGCATTTGC ACAGGGGAGG GGGGTGCCCT 720
 CCTTCCTTAG GACCTGGGGG CCAGGCTGAC TTGGGGGGCA GACTTGACAC TAGGCCCCAC 780
 TCACCTAGAT GTCTGAAAT TCCACCAGG GGGTCACCTT GGGGGTTTAG GGACCTATTT 840
 TTAACACTAG GGGCTGGCCC ACTAGGAGGG CTGGCCCTAA GATACAGACC CCCCCAACTC 900
 CCAAAGCGG GGAGGAGATA TTTATTTTGG GGAGAGTTTG GAGGGGAGGG AGAATTTATT 960
 10 AATAAAGAA TCTTTAACTT TAAAAAATA AAAAAAAA

Seq ID NO: 631 Protein sequence
Protein Accession #: NP_057723.1

15 1 11 21 31 41 51
 MARGSLRRLL RLLVLGLWLA LLRSVAGEQA PGTAPCSRGS SWSADLDKCM DCASCRARPH 60
 SDFCLGCAAA PPAPFRLLWP ILGGALSLTF VLGLLSGLFV WRCRRREKF TPIESTGGE 120
 20 GCPAVALIQ

Seq ID NO: 632 DNA sequence
Nucleic Acid Accession #: NM_003816.1
Coding sequence: 79..2538

25 1 11 21 31 41 51
 CGGCAGGGTT GGAATATGAT GGAAGAGGCG GAGGTGGAGG CGACCGAGTG CTGAGAGGAA 60
 CCTGCGGAAT CGCGCGAGAT GGGGTCTGGC GCGCGCTTTC CCTCGGGGAC CCTTCGTGTC 120
 CGGTGTGTGG TGTGTCTTGG CCTGGTGGGC CCAGTCTCTG GTGCGGCGCG GCCAGGCTTT 180
 30 CAACAGACCT CACATCTTTC TTCTTATGAA ATTATAACT CTGGAGATT AACTAGAGAA 240
 AGAAGAGAAG: CCCTCAGGCC CTATTCAAAA CAAGTATCTT ATGTTATTCA GGCTGAAGGA 300
 AAGAGCATA TTATTCACTT GGAAGGGAAC AAAGACCTTT TGCTGAAGA TTTGTGTTT 360
 TATACTTACA ACAAGGAAGG GACTTTAATC ACTGACCATC CCAATATACA GAATCATTTG 420
 CATTATCGGG GCTATGTGGA GGGAGTTTAT AATTCATCCA TTGCTCTTAG CGACTGTTTT 480
 35 GGACTCAGAG GATTGCTGCA TTTAGAGAAT GCGAGTTATG: GGATTGAACC CCTGCAGAAC 540
 AGCTCTCAT TTAGACACAT CATTATTCGA ATGGATGATG TCTACAAGA GCCTCTGAAA 600
 TGTGGAGTTT CCAACAAGGA TATAGAGAAA GAAACTGCAA AGGATGAAGA GGAAGAGCCT 660
 CCCAGCATGA CTGAGTACT TCGAAGAAGA AGAGCTGTCT TGCCACAGAC CCGGTATGTG 720
 GAGCTGTTCA TTGTCTAGA CAAGGAAAGG TATGACATGA TGGGAAGAAA TCAGACTGCT 780
 40 GTGAGAGAAG AGATGATCTT CCTGGCAAAC TACTTGGATA GTATGTATAT TATGTTAAAT 840
 ATTGCAATTG TGCTAGTTGG ACTGGAGATT TGGACCAATG GAAACCTGAT CAACATAGTT 900
 GGGGTGTGCT GTGATGTGCT GGGGAACCTC GTGCAGTGGC GGGAAAAGTT TCTTATCACA 960
 CGTCGAGAGC ATGACAGTGC ACAGCTAGTT CTAAGAAGAG GTTTTGGTGG AACTGCGAGG 1020
 45 ATGGCATTTG TGGGAACAGT GTGTTCAAGG AGCCACGCGG GCGGGATTAA TGTGTTTGA 1080
 CAAATCACTG TGGAGACATT TGCTTCCATT GTTCTCATG AATTTGGTCA TAATCTTGA 1140
 ATGAATCAGC ATGATGGGAG AGATTGTTCC TGTGGAGCAA AGAGCTGCAT CATGAATTCA 1200
 GGAGCATCGG GTTCCAGAAA CTTTAGCAGT TGCAGTGCAG AGGACTTTGA GAAGTTAACT 1260
 TTAATAAAGG GAGGAAACTG CCTTCTTAAT ATTCCAAGC CTGATGAAGC CTATAGTGCT 1320
 CCCTCTGTGG TGAATAAGTT GGTGGACGCT GGGGAAGAGT GTGACTGTGG TACTCCAAAG 1380
 50 GAATGTGAAT TGGACCTCTG CTGCGAAGGA AGTACCTGTA AGCTTAAATC ATTTGCTGAG 1440
 TGTGCATATG GTGACTGTTG TAAAGACTGT CGGTTCCCTC CAGGAGGTAC TTTATGCCGA 1500
 GGAAAAACCA GTGAGTGTGA TGTTCAGAG TACTGCAATG GTTCTTCTCA GTTCTGTGAG 1560
 CCAGATGTTT TTATTCAGAA TGGATATCCT TGCCAGAAAT ACAAGCCCTA TTGTCTAAC 1620
 55 GGATGTGCTG AGTATATGTA TGCTCAATGT CAAGTCACTT TTGGCTCAAA AGCCAAGGCT 1680
 GCCCCCAAAG ATTGTTTCAT TGAAGTGAAT TCTAAGGTG ACAGATTGGG CAATTGTGTT 1740
 TTCTCTGGCA ATGAATACAA GAAGTGTGCC ACTGGGAATG CTTTGTGTGG AAAGCTTCAG 1800
 TGTGAGAATG TACAAGAGAT ACCTGTATT TGGAAATGTG CTGCTATTAT TCAAAGCCTC 1860
 AGTTCAGGCA CCAATGTTTG GGGTGTGGAT TTCCAGCTAG GATCAGATGT TCCAGATCTC 1920
 60 GGGATGTTTA ACGAAGGCAC AAAATGTGGT GCTGGAAGA TCTGTAGAAA CTTCCAGTGT 1980
 GTAGATGCTT CTGTCTGAAA TTATGACTGT GATGTTTACA AAAAGTGTCA TGGACATGGG 2040
 GTATGTAAAT GCAATAAGAA TTGTCACTGT GAAATGSGCT GGGCTCCCCC AAATTGTGAG 2100
 ACTAAAAGAT ACGGAGGAAG TGTGGACAGT GGACCTACAT ACAATGAAT GAATAGTCCA 2160
 TTGAGGAGCG GACTCTCTGT CTCTCTCTC CTAATTGTTT CCCTTATTGT CTGTGCTATT 2220
 65 TTTATCTTCA TCAAGAGGGA TCAACTGTGG AGAAGCTACT TCAGAAAAGAA GAGATCACAA 2280
 ACATAGAGAT CAGATGGCAA AAATCAAGCA AACCTTCTA GACAGCCGGG GAGTGTCTCT 2340
 CGACATGTTT CTCCAGTGAC ACCTCCAGAA GAAGTTCCTA TATATGCAAA CAGATTGCA 2400
 GTACCAACCT ATGCAGCCAA GCAACCTCAG CAGTTCCCAT CAAGGCCACC TCCACCACAA 2460
 CCGAAAGTAT CATCTCAGGG AAACCTTAAT CCTGCCCTGC CTGCTCCTGC ACCTCCTTTA 2520
 70 TATAGTCCC TCACTGATT TTTTAACT TCTTTTGA AATGCTTCA GGGAACTGAG 2580
 CTAATACCTT TTTTCTTCT TGATGTTTC TTGAAAGCC TTTCTGTTG AACTATGAAT 2640
 GAAAAACAAA CACCAACAAA CAGACTTCAC TAACACAGAA AACAGAAAC TGAGTGTGAG 2700
 AGTTGTGAAA TACAAGGAAA TGCAGTAAAG CCAGGGAAIT TACAATAACA TTTCCGTTTC 2760
 CATCATGAAA TAAGTCTTAT TCAGTCACTG GTGAGGTTAA TGCACTAATC ATGGATTTT 2820
 75 TGAACATGTT ATTGCACTGA TTCTCAAAAT AACTGTATTG GTGTAAGATT TTTGTCAAT 2880
 AGTGTTTAAG TGTATTCTG AATTTCTAC CTTAGTTATC ATTAATGTAG TTCTCATTTG 2940
 AACATGTGAT AATCTAATAC CTGTGAAAAC TGACTAATCA GCTGCCAATA ATATCTAATA 3000
 TTTTTCATCA TGCAAGAAAT AATAATCATC ATACTCTAGA ATCTGTCTG TCACTCACTA 3060
 CATGAATAG CAAATATTGT CTTCAAAAGA ATGCACAAGA ACCACAATTA AGATGTCTA 3120
 80 TTAATTTGAA AGTAAATAAT ACTATAAAG AGTGTGTGTG TATTCACGCA GTTACTCGCT 3180
 TCCATTTTAA TGACCTTTCA ACTATAGSTA ATAATCTTCA GAGAAATTA TTTAATATTA 3240
 GAATTTCTAT TATGAATCAT GTGAAAGCAT GACATTCGTT CACAATAGCA CTATTTTAAA 3300
 TAAATTTATA GCTTTAAGGT ACGAAGTATT TAATAGATCT AATCAAAATAT GTTGATTACT 3360
 GGCATAATAA AAGCAGGAGC AATTATATAA TCTTCAATCA ATTGAACCTT TACAAAAACA 3420
 85 CTTGAGAATT TCATAGCAC TTTAAATCT GAACCTTCAA AGCTTGCTAT TAAATCAATT 3480
 AGAATGTTTA CATTTACTAA GGTGTGCTGG GTCATGTAAA ATATTAGACA CTAATATTTT 3540
 CATAGAAATT AGGCTGGAGA AAGAAGGAAG AAATGGTTTT CTAAATACC TACAAAAAAG 3600
 TTACTGTGCT ATCTATGAGT TATCATCTTA GCTGTGTTAA AAATGAATT TTAATATGGC 3660

AGATATGGTA TGGATCGTAA AATTTTAAGC ACTAAAAATT TTTTCATAAC CTTTCATAAT 3720
AAAGTTTAAAT AATAGGTTTA TTAAGTGAAT TTCATTAGTT TTTTAAAAAGT GTTTTGTGTT 3780
TGTGTATATA TACATATACA AATACAACAT TTACAATAAA TAAATAACTT GAAATTCCTCA 3840
AAAAAAAAA AAAAAAAAAA AAAAA

Seq ID NO: 633 Protein sequence
Protein Accession #: NP_003807.1

1 11 21 31 41 51
MGSGARPPSG TLRVRMLLLL GLVGPVLGAA RFGFQQTSHL SSYEIITPWR LTRERREAPR 60
PYSKQVSYVI QAEGKEHIIH LERNKDLLPE DFFVVTYNKE GTLITDHPNI QNHCHYRGYV 120
EGVHNSSIAL SDCFLRLGLL HLENASYGIE PLQNSSHFEH IYRMDVYK EPLKCGVSNK 180
DIEKETAKDE EEEPPSMTQL LRRRRVLPQ TRYVELFIVV DKERYDMMGR NQTAVREEMI 240
LLANYLDSMY IMLNIRIVLV GLEIWTNGL INIVGAGDV LGNFVQWREK FLITRRRHS 300
AQLVLKKGFG GTAGMAFVGT VCSRSAGGI NVFGQITVET FASIVAEHLG HNLGMNHDDG 360
RDSCGAKSC IMNSGASGR NFSSCSAEDF EKLTNLNKGK CLNIPKPEE AYSAPSCGNK 420
LVDAGEBCDC GTPKECELDP CCEGSTCKLK SFAECAYGDC CKDCRFLPGS TLCRGKTSEC 480
DVPEYCNSS QFCQPVFIQ NGYPCQNKKA YCYNGMCQYV DAQCQVIFGS KAKAAPKDCF 540
IEVNSKQDRF GNCCFSGNEY KKCATGNALC GKLCQENVQE IPVFGIIPAI IQTPSRGKTC 600
WGVDFQLGSD VPDPMVNEG TKCGAGKICR NFQCVDAVL NYDCDVQKCC HGHGVCSNKK 660
NCHCENGWAP PNCETKGYGG SVDGPTYN MNTALRDGLL VFFFLIVPLI VCAIFIFIKR 720
DQLWRSYFRK KRSQTYESDG KNQANPSRQP GSVPRHVSVP TPPREVPIYA NRVAVPTYAA 780
KQPQFPSPRP PPPQPKVSSQ GNLIIPARPAP APPLYSSLT

Seq ID NO: 634 DNA sequence
Nucleic Acid Accession #: NM_002091.1
Coding sequence: 56..503

1 11 21 31 41 51
AGTCTCTGCT CTTCCAGCC TCTCCGCGC GCTCCAAGGS CTTCCTGCG GGACCATGCG 60
CGCGAGTAG CTCCGCTGG TCCTGCTGGC GCTGGTCTC TGCCATAGCG CCGGGGGCG 120
AGCGGTCCCG CTGCTGCGG GCGGAGGGAC CGTGCTGACC AAGATGTACC CGCGCGGCAA 180
CCACTGGGCG GTGGGCACT TAATGGGGAA AAAGAGCACA GGGGAGTCTT CTTCTGTTT 240
TGAGAGAGGG AGCCTGAAGC AGCAGCTGAG AGAGTACATC AGGTGGGAAG AAGCTGCAAG 300
GAATTTGCTG GGTCTCATAG AAGCAAAGGA GAACAGAAAC CACCAGCCAC CTCACCCCAA 360
GGCCTTGGCG AATCAGCAGC CTTCTGCGGA TTCAGAGGAT AGCAGCAACT TCAAAGATGT 420
AGGTTCAAAA GGCAAGTTG GTAGACTCTC TGCTCCAGGT TCTCAACGTG AAGGAAGGAA 480
CCCCGAGCTG AACCAAGCAAT GATAATGATG GCCTCTCTCA AAAGAGAAAA ACAAAACCCC 540
TAAGAGACTG AGTTCTGCAAT GCATCAGTTC TACGGATCAT CAACAAGATT TCCTTGTGCA 600
AAATATTGTA CTATTCTGTA TCTTTCATCC TTGACTAAAT TCGTGATTTT CAAGCAGCAT 660
CTTCTGGGTT AAACCTGTGTT GCTGTGAACA ATTGTGAAAA AGAGTCTTCC AATTAATGCT 720
TTTTTATATC TAGGCTACCT GTTGGTTAGA TTCAAGGCCC CGAGCTGTGA CCATTACAAA 780
TAAAGCTTA AACACAT

Seq ID NO: 635 Protein sequence
Protein Accession #: NP_002082.1

1 11 21 31 41 51
MRGSELPLVL LALVLCAPR GRAVFLPAGG GTVLTMYPR GNHWAUGHLM GKKTGESSS 60
VSEKSLKQKQ LREYIRWEA ARNLLGLIEA KENRNHQPQP PKALGNQPPS WDESDSNFK 120
DVSGSKGVGR LSAPGSQREG RNPQLNQ

Seq ID NO: 636 DNA sequence
Nucleic Acid Accession #: NM_016522.1
Coding sequence: 265..1299

1 11 21 31 41 51
GCGGAAGCAG CGAGGAGGGA GCCCCTTTG GCCGTCTCC GTGGAACCGG TTTTCCGAGG 60
CTGGCAAAAG CCGAGGCTGG ATTGGGGGA GGAATATTAG ACTCGGAGGA GTCTGCGCGC 120
TTTTCTCTC CCGCGCTCTC CCGGTGCGCG CGGGTTCAAC GCTCAGTCCC CGCGCTCGCT 180
CCGCACCCCA CCACTCTCTC GTGCTGCGCC GGGGGGCGTG TGCCGTGCGG CTGCGGAGT 240
TCGGGGAAGT TGTGGCTGTC GAGAATGGGG GTCTGTGGGT ACCTGTTCCT GCCCTGGAAG 300
TGCTCTGTCG TCGTGTCTCT CAGGCTGCTG TTCCTGTGAC CCACAGAGT GCCCGTGCOC 360
AGCGGAGATG CCACTCTCCC CAAAGCTATG GACAACGTGA CGGTCCGGCA GGGGGAGAGC 420
GCCACCTCA GGTGCACTAT TGACAACCGG GTCAACCGGG TGGCCTGGCT AAACCGCAGC 480
ACCATCTCTC ATGCTGGGAA TGACAAGTGG TGCTGTGATC CTGCGTGGT CTTCTGAGC 540
AACACCCAAA CGCAGTACAG CATCGAGATC CAGAAGCTGG ATGTGTATGA CGAGGGCCCT 600
TACACCTGCT CGGTGAGAGC AGACAACAC CCAAGACCT CTAGGGTCCA CCTCATTGTG 660
CAAGTATCTC CCAAAATTTG AGAGATTCTC TCAGATATCT CCATTAATGA AGGGAACAAT 720
ATTAGCTCA CCGTCAATAG AACTGGTAGA CCAGAGCCTA CGGTACTTGT GAGACATC 780
TCTCCCAAAG CGGTTGGCTT TGTGAGTGAA GACGAATACT TGGAAATTCA GGGCATCACC 840
CGGGAACAGT CAGGGGACTA CGAGTGCAGT GCCTCCAATG ACGTGGCCGC GCCCGTGGTA 900
CGGAGAGTAA AGGTCAACGT GAACATATCA CCATACATTT CAGAAGCCAA GGGTACAGGT 960
GTCCCGTGG GACAAAGGG GACACTGCAG TGTGAAGCCT CAGCAGTCCC CTCAGCAGAA 1020
TTCCAGTGTG ACAAGATGA CAAAGACTG ATTGAAGGAA AGAAAGGGGT GAAAGTGGAA 1080
AACAGACCTT TCCTCTCAAA ACTCATCTTC TTCATGTCT CTGAACATGA CTATGGGAAC 1140
TACATTGCG TGCCCTCCAA CAAGCTGGGC CACACCAATG CCAGCATCAT GCTATTGTGT 1200
CCAGGCGCGC TCAGCGAGGT GAGCAACGGC ACGTCCAGGA GGGCAGGCTG CGTCTGGCTG 1260
CTGCTCTCT TGGTCTTGCA CCTGCTTCTC AAATTTTGAT GTGAGTGGCA CTTCCCAACC 1320
CGGGAAGGCG TCGCGCACCC ACCACCAACA ACACAACAGC AATGGCAACA CCGCAGCAGAA 1380
CCAATCAGAT ATATACAAAT GAAATTAGAA GAAACACAGC CTCATGGGAC AGAAATTGTA 1440
GGGAGGGGAA CAAAGATATC TTTGGGGGGA AAGAGGTTTT AAAAAAGAAA TTGAAAATTG 1500
CCTTGAGAT ATTTAGGTAC AATGGAGTTT TCTTTTCCCA AACGGGAAGA ACACAGCACA 1560

WO 02/086443

CCGCGCTTGG ACCACTGCA AGCTGCATCG TGCAACCTCT TTGGTGCCAG TGTGGGCAAG 1620
 GGCTCAGCCT CTCTGCCCC AGACTGCCCC CACGTGGAAC ATTCTGGAGC TGGCCATCCC 1680
 AAATTCATTC AGTCCATAGA GACGAACAGA ATGAGACCTT CGGCCCCAAG CGTGGCGCCT 1740
 CCGGCCCAAG CGTGGCGCTG CGGGCACTTT GGTAGACTGT GCCACCAAG CGTGTGTGT 1800
 GAAACGTGAA ATAAAAAGAG CAAAAAATA AAAAAAATA

Seq ID NO: 637 Protein sequence
 Protein Accession #: NP_057606.1

1 11 21 31 41 51
 MGVCGLPLP WKCLVVSILR LLFLVPTGVP VRSGDATFPK AMDNVTVRQG ESATLRCTID 60
 NRVTTRVAVLN RSTILYAGND KWCLDPRVVL LSNTQTQYSI EIQNVVDVDE GPYTCSVQTD 120
 NHPKTSRVHL IVQVSPKIVE ISSDISINEG NNISLTCTAT GRPEPTVTWR HISPKAVGFV 180
 SEDEYLEIQG ITRQSGDYVE CSASNDVAAP VVRVKVTVN YPPYISEAKG TGVVPVGKGT 240
 LQCCASAVPS AEFQWYKDDK RLIEGKKGVK VNRPFPLSKL IFFNVSEHDY GNYTCVASNK 300
 LGHTNASIML FGPVAVSEVS NGTSRRAGCV WLLPLLVHL LKLF

Seq ID NO: 638 DNA sequence
 Nucleic Acid Accession #: NM_012261.1
 Coding sequence: 203..1045

1 11 21 31 41 51
 GATTGTCTCT GCCAGCAGCT GTCGGTGCCG CGCTGACAC CGAGTCTTAG CTAGGCGCTC 60
 ACAGAATACG CGCTCCCTCC CTCGCCCTTC TCTGTCCCCC GCCTCTGCT CAACCCGGGC 120
 CACTCCAGCG CGGACTTTGA GGGATTCCCT CTCTGGCGGC CTCTGCAGCA GCACAGCCGG 180
 CCTCATTCGG GGCACGTGCA GTATGGATCT CCAAGGAAGA GGGTCCCCA GCATCGACAG 240
 ACTTCGAGTT CTCTGATGT TGTTCATAC AATGGCTCAA ATCATGGCAG AACAGAAGT 300
 GGAATAATCTC TCAGGCTTTT CCACTAACCC TGAATAAGAT ATATTGTGG TCGGGGAAA 360
 TGGGACGACG TGTCTCATGG CAGAGTTTGC AGCCAAATTT ATTGTACCTT ATGATGTGTG 420
 GGCCAGCAAC TACGTATGAT TGATCAGAGA ACAGGCGGAT ATCGCATTGA CCGGGGAGC 480
 TGAGGTGAGG GGGCGCTGTG GCCACAGCCA GTCCGAGCTG CAAGTGTCTT GGGTGGATCG 540
 CGCATATGAG CTCAAATGCG TCTTTGTAAA GGAAGGCCAC AACATGTCCA AGGGACCTGA 600
 GGGCACTTGG AGGCTGAGCA AAGTGCAGTT TGTCTACGAC TCCTCGGAGA AAACCCACTT 660
 CAAAGACGCA GTCAGTGTCT GGAAGCACAC AGCCAACTCG CACCACCTCT CTGCCTTGGT 720
 CACCCCGCTT GGAAGTCTCT ATGAGTGTCA AGCTCAACAA ACCATTTCAC TGGCCTCTAG 780
 TGATCCGAGC AAGACGCTCA CCATGATCCT GTCTCGGCTC CACATCCAAC CTTTTCACAT 840
 TATCTCAGAT TTTGTCTTCA GTGAAGAGCA TAAATGCCCA GTGGATGAGC GGGAGCAACT 900
 GGAAGAAACC TTGCCCCGTA TTTGGGGCT CATCTGGGCG CTGCTATCA TGGTAACACT 960
 CGCGATTAC CAGTCCACCC ACAAAATGAC TGCCAACCAAG GTGCAGATCC CTCGGGACAG 1020
 ATCCAGTAT AAGCAGATGG GCTAGAGGCC GTTAGGCAGG CACCCCTAT TCCTGCTCCC 1080
 CCAACTGGAT CAGGTAGAAC AACAAAGCA CTTTCCATC TTGTACACGA GATACACAA 1140
 CATAGCTACA ATCAACAGG CTTGGGTATC TGAGGCTTGC TTGGCTTGTG TCCATGCTTA 1200
 AACCCACGGA AGGGGGAGAC TCTTTCGGAT TTGTAGGGTG AAATGGCAAT TATTCTCTCC 1260
 ATGCTGGGGA GGAAGGGAGG AGGGTCTCAG ACAGCTTTCG TGCTCATGTT GGCCTTGGCT 1320
 TGACTCTCCA AAGAGCAATA AATGCCACTT GGAGCTGTAT CTGGCCCCAA AGTTTAGGGA 1380
 TTGAAACAT GCTTCTTGA GGAGGAAACC CCTTATGTTT CAGAAGAATA TGGGGTGCTT 1440
 TGCTCCCTTG GACACAGCTG GCTTATCCTA TACAGTTGTC AATGCACACA GAATACAAAC 1500
 TCATGCTCCC TGCAGCAAGA CCCCTGAAAG TGATTATGTC TTCTGGCTGG CATTCTGCAT 1560
 GTTTAGTGAT TGTCTTGGGA ATGTTTCACT GCTACCCGCA TCCAGCGACT GCAGCACCAG 1620
 AAAACGACTA ATGTAACATAT GCAGAGTTGT TTGGACTTCT TCCTGTGCCA GGTCCAAAGT 1680
 GGGGACCTG AAGAATCAAT CTGTGTGAGT CTGTTTTTCA AAATGAAATA AAACACACTA 1740
 TTCTCTGGC

Seq ID NO: 639 Protein sequence
 Protein Accession #: NP_036393.1

1 11 21 31 41 51
 MDLQGRGVPS IDRLRLVLM L FHTMAQIMAB QEVENLSGLS TNPEKDIFV RENGITCLMA 60
 EFAAKPIVPI DVWASNYVDL ITEQADIALT RGAEVKGRCG HSQSEIQVFN VDRAYALKML 120
 FVKESHNMKS GPEATWRLSK VQFVYDSSEK THFKDAVSAG KHTANSHLS ALVTPAGKSY 180
 ECQAQQTISL ASSDPQKTVT MILSAVHIQF FDIISDFVFS EEHKCPVDER EQLEETLPLI 240
 LGLILGLVIM VTLAIYHVHH KMTANQVQIP RDRSQYKHM

Seq ID NO: 640 DNA sequence
 Nucleic Acid Accession #: NM_002993.1
 Coding sequence: 64..408

1 11 21 31 41 51
 GGCAGAGCC AGTCTCCGGC CTCCACCCA GCTCAGGAAC CCGCGAACC TCTCTGACC 60
 ACTATGAGC TCCCGTCCAG CGCGCGGCGG CGTGTCCGGG GTCCTTCGGG TCCTTGTGC 120
 GCGTGTCTCG CGCTGCTGCT CCGTGTGACG CGCGCGGGG CCCTCGCCAG CGCTGGTCT 180
 GTCTCTGCTG TGCTGACAGA GCTGCGTTCG ACTTGTATTAC GCGTTACGCT GAGAGTAAAC 240
 CCAAAACGA TTGGTAAATC GCAGGTGTTT CCGCAGGCC CGCAGTGTCT CAAGTGGAA 300
 GTGGTAGCCT CCCTGAAGAA CGGGAAGCAA GTTGTCTGG ACCCGGAAGC CCCTTTCTA 360
 AAGAAAGTCA TCCAGAAAT TTTGGACAGT GGAAACAAGA AAACTGAGT AACAAAAAG 420
 ACCATGCATC ATAAAAATGC CAGTCTTCA GCGAGCAGT TTTCTGGAGA TCCCTGACC 480
 CAGTAAGAA TAAAGGAAG GGTGGTTTT TTTCCATTT CTACATGGAT TCCCTACTT 540
 GAAGAGTGTG GGGGAAAGCC TACGCTTCT CCGTGAAGTT ACAGCTCAGC TAATGAAGTA 600
 CTAATATAGT ATTTCCACTA TTTACTGTTA TTTTACCTGA TAAGTTATTG AACCCCTTGG 660
 CAATTGACCA TTTGTGAGC AAGAATCAC TGGTTATTAG TCTTCAATG AATATTGAAT 720
 TGAAGATAAC TATTGTATT CTATCATACA TTCCTTAAAG TCTTACCGAA AAGGCTGTGG 780
 ATTTCGTATG GAAATAATGT TTTATTAGT TGCTGTTGAG GGAGGTATCC TGTGTCTT 840
 ACTCACTCT CTCATAAAT AGGAAATATT TTAGTCTGT TTTCTTGGG AATATGTTAC 900

TCTTTACCTT AGGATGCTAT TTAAGTTGTA CTGTATTAGA AACTTGGGTG TGTCAATCCG 960
 TTATCTGTGC AGAATATATT TCCTTATTCA GAATTTCTAA AAATTTAAGT TCTGTAAAGG 1020
 CTAATATATT CTCTTCTAT GGTTTTAGAT GTTTGTAGTC TTCCTAGTAT GGCATAATGT 1080
 CATGATTTAC TCATTAAACT TTGATTTTGT ATGCTATTTT TTCACTATAG GATGACTATA 1140
 ATTCTGGTCA CTAATATAC ACTTTAGATA GATGAAGAAG CCAAAAACA GATAAATTC 1200
 TGATTGCTAA TTTACATAGA AATGTATTCT CTGTGTTTTT TAAATAAAG CAAAATTAAC 1260
 AATGATCTGT GCTCTGCAAA GTTTTGAAAA TATATTTGAA CAATTTGAAT ATAAATTCAT 1320
 CATTTAGTCC TCAAAATATA TACAGCATTG CTAAGATTTT CAGATATCTA TTGTGGATCT 1380
 TTTAAAGGTT TTGACCATTT TGTATGAGG AATTATACAT GTATCACATT CACTATATTA 1440
 AAAATGCACT TTTATTTTTT CCTGTGTGTC ATGTTGGTTT TTGTTACTTG TATTGTCAAT 1500
 TGGAGAAACA ATAAAGATT TCTAAACCAA AAAAAA AAAAAA

Seq ID NO: 641 Protein sequence
 Protein Accession #: NP_002984.1

1 11 21 31 41 51
 MSLPSSRAAR VPGPSGLCA LLALLLLTLP PGPLASAGPV SAVLTELRCT CLRVTLRVNP 60
 KTIGKLQVFP AGPQCKSKEV VASLKNKGQV CLDPEAPFLK KVIQKILDSG NKKK

Seq ID NO: 642 DNA sequence
 Nucleic Acid Accession #: NM_013271.1
 Coding sequence: 27..809

1 11 21 31 41 51
 TCCGAGGCA GGCCTGCTGG GGCAGCATGG CCGGCTGCGC GCTGCTCTGG GGGCCGCGGG 60
 CCGGGGGCGT GCGCCTTTTG GTGCTGTGTC TGCTCGGCCT GTTTCGGCCG CCCCCCGCGC 120
 TCTGCGCGCG GCGCGTAAAG GAACCCCGCG GCCTAAGCGC AGCGTCTCCG CCGTTGGCTG 180
 AGACTGGCGC TCCTGCGCGC TTCGCGCGGT CAGTGCCCGG AGGTGAGGCG GCGGGGCGCG 240
 TGCAGGAGCT GCGCGCGCGC CTGCGCGCAT TGCTGGAGGC CGAACGTCAG GAGCGGGCGC 300
 GGGCCGAGGC GCAGGAGGCT GAGGATCAGC AGGCGCGCGT CCGGCGCAG CTGCTGCGCG 360
 TCTGGGCGCG CCCCCGCAAC TCTGATCCGG CTCTGGGCGT GAGCGACGAC CCGCAGCGCG 420
 CTGCGAGCGA GCTCGCTGCG GCTCTGCTCC GCGCCCGCGT GCGCCGCGCT TGACCTAGCG CCGCTAGCAG 480
 CCCAGCTTGT CCGCGCGCGC GTCCCCCGCG CCGCGCTCCG ACCCGCGCCC CCGGTCTACG 540
 ACAGCGCGCC CCGCGCGCGC GATGCTGAGG AGGCAGGCGA CGAGACACCC GACGTGAGCC 600
 CCGAGCTGTT GAGGTACTTG CTGGGACGGA TTCTTGCGGG AAGCGCGGAC TCGAGGGGG 660
 TGGCAGCCCC GCGCGCGCTC CGCGTGCGG CCGACCCAGA TGTGGGCTCT GAGCTGCCCC 720
 CTGAGGGCGT GCTGGGGGCG CTGCTGCGTG TGAACCGCT AGAGACCCCG GCGCCCGAG 780
 TGCCCTGCAG CCGCTCTTTG CCACCTGAG CACTGCGCGG ATCCCGTGCA CCGTGGGACC 840
 CAGAAGTGCC CCGCCATCC CCGCACCAGG ACTTCTCCCC GCCAGCACGT CCAGAGCAAC 900
 TTACCCCGCG CAGCCAGCCC TCTCACCCGA GGATCCCTAC CCGCTGGCCC ACAATAACAT 960
 GATCTGAGC

Seq ID NO: 643 Protein sequence
 Protein Accession #: NP_037403.1

1 11 21 31 41 51
 MAGSPLLWPF RAGGVGLLVL LLLGLFRPPP ALCARPVKEP RGLSASAPPL AETGAPRRFR 60
 RSVPRGEAAG AVQELARALA HLLAEARQER ARAEAQEAED QQARVLAQLL RVWGAPRNSD 120
 PALGLDDDPD APAALARAL LRARLDPAAL AAQLVPAPVP AAALRPRPPV YDDGPAFEDA 180
 EEAGDETPDV DEELLRYLLG RILAGSADSE GVAAPRRLRR AADHDVGSSEL PPGVGLGALL 240
 RVKRLETAP QVPARRLLPP

Seq ID NO: 644 DNA sequence
 Nucleic Acid Accession #: NM_002214
 Coding sequence: 681..2990

1 11 21 31 41 51
 CCCAGAGCGC CTTCCCGCTG TTGCTGGCAT CCGGAGCTTC CTCCCTTGCC AGCCAGGAGC 60
 CTGCGGACTT GTCTTTGGCC GCTGCTCCGC AGACGGGGCT GCAAAGCTGC AACTAATGGT 120
 GTTGGCCCTC CTGCCCCACT GTGGAAGCAA CTGCGCTGAT TGATGCGCCA CAGACTTTTT 180
 TCCCTCGAC CTCGCGCGCG TACCTCCCA CAGATCCAGC ATCACCAGT GAATGTACAT 240
 TAGGTTGGTT TCCCGCCAG CTTCGGGCTT TGTTTGGGTT TGATTGTGTT TGGCTCTTCG 300
 CTAAGCTGAT TTATGACAGA GAAGCCCCAC CCGCTGGAGA GAAACAAAAG CTCTTTTCTT 360
 TGTCCCGGAG CAGGCTGCGG AGCCCTTGCA GAGCCCTCTC TCCAGTCCGC GCGCGGCCCT 420
 TGGCGCTCGA AGGAGGTGCT TCTCGCGGAG ACCGCGGGAC CCGCGTGCC GAGCCGGGAG 480
 GGCGTAGGG GCGCTGAGAT GCGGAGCGGT GCGCGGGCCC GCTTACCTGC ACCGCTTGCT 540
 CCGAGCGCG GGGTCCGCT GCTAGGCTTG CGGAAACGT CCTAGCGACA CTCGCGCGCG 600
 GCGCGCGAG TCGCGCGGGA GCGCGAGCCC GCGTCCGGAA GGCAGCCAGG CCGCGGGCGC 660
 GGGCGGGCT GTTTTGCAAT ATGTGCGGCT CCGCCCTGGC TTTTITTAAC GCTGCAATTG 720
 TCTGCTGCA AAACGACCGG CAGGTTCCCG CTCTGTTCT CTGGGAGGCC TGGGTGTTTT 780
 CACTTGTCT TGGACTGGGC CAGGTGAAG ACAATAGATG TGCACTTCA AATGACGAT 840
 CTTGTGCCAG GTGCTTGGC CTGGGTCAG AATGTGATG GTGTGTTCA AGGATTTC 900
 TTTAGGTGG ATCAAGAGT GAACGTTGTG ATATTGTTTC CAATTTAATA AGCAAAGGCT 960
 GCTCAGTTGA TTCAATAGAA TACCCATCTG TGCACTTTAT AATACCCACT GAAATGAAA 1020
 TTAATACCCA GGTGACACCA GGAGAAGTGT CTATCCAGCT GGTCCAGGA GCCGAAGCTA 1080
 ATTTATGCT GAAAGTTCAT CCTCTGAAGA AATATCTGTT GGATCTTTAT TATCTTGTG 1140
 ATGTCTCAGC ATCAATGAC AATAATATAG AAAAATTAAT TTCCGTTGGA AACGATTTAT 1200
 CTAGAAAAT GGCAATTTTC TCCCGTACT TTGCTCTGG ATTGGCTCA TACGTTGATA 1260
 AAACAGTTTC ACCATACATT AGCATCCACC CCGAAGGAT TCAATATCAA TGCAGTGACT 1320
 ACAATTTAGA CTGCATGCT CCCCATGGAT ACATCCATGT GCTGCTTTG ACAGAGAACA 1380
 TCACTGAGTT TGAGAAAGCA GTTCATAGAC AGAAGATCTC TGGAAACATA GATACACCAG 1440
 AAGAGGTTT TGACGCCATG CTTCAGGCAG CTGTCTGTGA AAGTCATATC GGATGGCGAA 1500
 AAGAGGCTAA AAGATTGCTG CTGGTGATGA CAGATCAGAC GTCTCATCTC GCTCTTGATA 1560

5	GCAAATTTGGC	AGGCATAGTG	GTGCCCAATG	ACGGAAACTG	TCATCTGAAA	AACAAGCTCT	1620
	ACGTCAAATC	GACCAACCATG	GAACACCCCT	CACTAGGCCA	ACTTTCAGAG	AAATTAATAG	1680
	ACAACAACAT	TAATGTGATC	TTTGCACTTC	AAGGAAACCA	ATTTTCATTG	TATAAGGATC	1740
	TTCTACCCCT	CTTGCCAGGC	ACCATTGCTG	GTGAAATAGA	ATCAAAGGCT	GCAACCTCA	1800
	ATAATTTGGT	AGTGAAGGCC	TATCAGAAGC	TCATTTCCAG	AGTGAAGTGT	CAGGTGGAAG	1860
	ACCAGGTACA	AGGCATCTAT	TTTAACATTA	CGGCCATCTG	TCCAGATGGG	TCCAGAAAGC	1920
	CAGGCATGGA	AGGCATGAGA	AACGTGACGA	GCAATGATGA	AGTCTTTTTC	AATGTAAACAG	1980
	TTACAATGAA	AAAATGTGAT	GTCAACAGAG	GAAGAACTA	TGCAATAATC	AAACCTATTG	2040
10	GTTTTAATGA	AACCGCTAAA	ATTCAATATC	ACAGAACTG	CAGCTGTGAG	TGTGAGGACA	2100
	ACAGAGGACC	TAAAGGAAAG	TGTGTAGATG	AAACTTTTCT	AGATTCCAAG	TGTTTCCAGT	2160
	GTGATGAGAA	TAAATGTGAT	TTTGATGAAG	ATCAGTTTTC	TTCTGAGAGT	TGCAAGTCAC	2220
	ACAAGGATCA	GCCTGTTTGC	AGTGGTGGAG	GAGTTTGTGT	TTGTGGGAAA	TGTTTCATGTC	2280
	ACAAATTTAA	GCTTGGAAAA	GTGTATGGAA	AATACTGTGA	AAAGGATGAC	TTTTCTTGTC	2340
	CATATCACC	TGGAATCTG	TGTGCTGGGC	ATGGAGAGTG	TGAAGCAGGC	AGATGCCAAT	2400
15	GCTTCAGTGG	CTGGGAAGGT	GATCGATGCC	AGTGCCCTTC	AGCAGCAGCC	CAGCACTGTG	2460
	TCAATTCAAA	GGGCCAAGTG	TGCAGTGGAA	GAGGCACGTG	TGTGTGTGGA	AGGTGTGAGT	2520
	GCACCGATCC	CAGGAGCATCT	GGCCGCTTCT	GTGAACACTG	CCCCACCTGT	TATACAGCCT	2580
	GCAAGGAAAA	CTGGAATTGT	ATGCAATGCC	TTCACTCTCA	CAATTTGTCT	CAGGTATATC	2640
20	TTGATCAGTG	CAAAACCTCA	TGTGCTCTCA	TGGAACAACA	GCAATTATGTC	GACCAAACTT	2700
	CAGAATGTTT	CTCCAGCCCC	AGCTACTTGA	GAATATTTT	CATCATTTTC	ATAGTTACAT	2760
	TCTTGATTGG	GTGCTTAA	GTCCTGATCA	TTAGACAGGT	GATACTACAA	TGGAATAGTA	2820
	ATAAAATTTAA	GTCTCATCA	GATTACAGAG	TGTGAGCCTC	AAAAAAGGAT	AAGTTGATTC	2880
	TGCAAAAGTG	TTGCAAGAA	GCAGTCACT	ACCGACGTGA	GAAGCCTGAA	GAATAAAAA	2940
25	TGGATATCAG	CAAAATTAAT	GCTCATGAAA	CTTTCAGGTG	CAACTTCTAA	AAAAAGATTT	3000
	TTAAACACTT	AATGGGAAAC	TGGAATTTGT	AATAATTTGT	CCTAAAGATT	ATAATTTTAA	3060
	AAGTCACAGG	AGGAGACAAA	TTGCTCACGG	TCATGCCAGT	TGCTGGTGTG	ACACTCGAAC	3120
	GAAGACTGAC	AAGTATCTCT	ATCATGATGT	GACTCACATA	GCTGCTGACT	TTTTTCAGAGA	3180
	AAAATGTGTC	TTACTACTGT	TTGAGACTAG	TGTGTTGTA	GCACTTTACT	GTAATATATA	3240
30	ACTTATTTAG	ATCAGCATAG	AATGTAGATC	CTCTGAAGAG	CACTGATTAC	ACTTTACAGG	3300
	TACCTGTTAT	CCCTACGCTT	CCCAGAGAGA	ACAATGCTGT	GAGAGAGTTT	AGCATTGTGT	3360
	CACATACAGG	GTACAGTAAT	CCCTGCACGT	GACATGTGAG	GAAGAAAAATA	ATCTGGCAAG	3420
	TATATTTCAA	GGTTGCCAAA	CACITCAACA	GTGTTGGTGT	GAATAGACAA	GAACAGCTAG	3480
	ATGAATAAAT	GATTTCGTGT	TCACITTTTC	AAGAGGTGAA	CAGATACAAC	CTTAATCTTA	3540
35	AAAGATTATT	GCTTTTAA	GTGTGTAGTT	TTATGCAATG	GTGTTTATGG	TTTGCTTATT	3600
	TTTGCAAGAT	GGATACTAAT	TCCAGCATTC	TCTCTCTTT	GCCITTTATGT	TTTGTTTTCT	3660
	TTTTTACAGG	ATAAGTTTAT	GTATGTCACA	GATGACTGGA	TTAATTAAGT	GCTAAGTTAC	3720
	TACTGCCATA	AAAACTAAT	AATACAATGT	CACITTTATCA	GAATACTAGT	TTTAAAGCT	3780
	GAATGTTAA						

Seq ID NO: 645 Protein sequence
Protein Accession #: NP_002205

45	1	11	21	31	41	51	
	MCQSALAFPT	AAFVCLQNDR	RGPASFLWAA	WVFSVLVLGL	QGEDNRCASS	NAASCARCLA	60
	LGPECGWCVQ	EDFTSGGSR	ERCDIVSNLI	SKGCSVDSIE	YPSVHVLIPT	ENEINTQVTP	120
	GEVSIQLRPG	AEANFMLKVH	PLKKYPVDLY	YLVDSVSMH	NNIEKLNSVG	NDLSRKMAPP	180
	SRDFRLGFS	YVDKTVSPY	SIHPERIHNQ	CSVDNLDMP	PHGYIHVLSL	TENITEPEKA	240
50	VHRQKISGNI	DTPGGGDFAM	LQAAVCESHI	GWRKEAKRLL	LVMTDQTSHL	ALDSKLAGIV	300
	VPNDGNCHLK	NNVYVKSTTM	EHPSLGLQSE	KLIDNNINVI	FAVQKQFHW	YKDLLPLLP	360
	TIAGEIESKA	ANLNNLVVEA	YQKLISEVKV	QVENQVQGIY	FNITAIKCPD	SRKPGMGCCR	420
	NVTSDNEVLF	NVTYTMKKCH	VTGKKNYAI	KPIGFNETAK	IHIHRNCSQ	CEDNRGPKGK	480
	CVDETFLDSK	CFQCDENKCH	FDEDQFSSES	CKSHKQDQVC	SGRGVVCVCK	CSCHKIKLKG	540
55	VYKYCEKDD	FSPCYHGNL	CAGHGECEAG	RQCFPSGWE	DRQCQPSAAA	QHCVNSKQV	600
	CSGRGTCVCG	RCECTDPRSI	GRFCEHCPTC	YTACKENWNC	MQCLHFNLS	QAILDQCKTS	660
	CLAMEQGHVY	DQTECEPSSP	SYLRIFFIIF	IVTFLIGLLK	VLIIRQVILQ	WNSNKKIKSS	720
	DYRVSASKKD	KLILQSVCTR	AVTYRREKPE	EIKMDISKLN	AHETFRCNF		

Seq ID NO: 646 DNA sequence
Nucleic Acid Accession #: NM_003318.1
Coding sequence: 1..2574

65	1	11	21	31	41	51	
	ATGGAATCCG	AGGATTTAAG	TGGCAGAGAA	TTGACAAATTG	ATTCCATAAT	GAACAAAGTG	60
	AGAGACATTA	AAAATAAGTT	TAAAAATGAA	GACCTTACTG	ATGAACTAAG	CTTGAATAAA	120
	ATTTTCTGCT	ATACTACAGA	TAACTCGGGA	ACTGTTAAAC	AAATTATGAT	GATGGCAAAC	180
	AACCCAGAGG	ACTGGTTGAG	TTTGTGCTC	AAACTAGAGA	AAACAGTGT	TCCGCTAAGT	240
70	GATGCTCTTT	TAAATAAAAT	GATTGGTCGT	TACAGTCAAG	CAATTGAAGC	GCTTCCCCCA	300
	GATAAATATG	GCCAAAATGA	GAGTTTGTCT	AGAATTCAAG	TGAGATTTCG	TGAATTAATA	360
	GCTATTCAAG	AGCCAGATGA	TGCACGTGAC	TACTTTCAAA	TGGCCAGAGC	AAACTGCAAG	420
	AAATTTGCTT	TGTTTCATAT	ATCTTTTGCA	CAATTTGAAC	TGTCACAAGG	TAAATGTCAA	480
	AAAAGTAAAC	AACTTCTTCA	AAAAGCTGTA	GAACGTGGAG	CAGTACCACT	AGAAATGCTG	540
75	GAAATTTGCC	TGCGGAATTT	AAACCTCCAA	AAAAAGCAGC	TGCTTTTCAG	GGAGGAAAAG	600
	AAGAATTTAT	CAGCATCTAC	GGTATTAACT	GCCCAAGAA	CATTTTCCGG	TTCACTTGGG	660
	CATTTACAGA	ATAGGAACAA	CAGTTGTGAT	TCCAGAGGAC	AGACTACTAA	AGCCAGGTTT	720
	TTATATGGAG	AGACATGCC	ACCACAAGAT	GCAGAAATAG	GTTACCGGAA	TTCAATTGAGA	780
	CAAACTAACA	AACTAAACAA	GTCAATGCCA	TTTGAAGAG	TCCAGITTA	CCTTCTAAAT	840
80	AGCCAGATT	GTGATGTGAA	GACAGATGAT	TCAGTTGTAC	CTTGTTTTAT	GAAGAGACAA	900
	ACCTCTAGAT	CAGAATCGCG	AGATTGGTGT	GTGCTGGAT	CTAAACCAAG	TGGAATATGAT	960
	TCCTGTGAAT	TAAAGAAATTT	AAAGTCTGTT	CAAAATAGTC	ATTTCAAGGA	ACCTCTGGTG	1020
	TCAGATGAAA	AGAGTTCTGA	ACTTATTATT	ACTGATTCAA	TAAACCTGAA	GAATAAAACG	1080
	GAATCAAGTC	TTCTAGCTAA	ATTAGAAGAA	ACTAAAGAGT	ATCAAGAAC	AGAGGTTCCA	1140
85	GAGAGTAACC	AGAAACAGTG	GCAATCTAAG	AGAAAGTCAG	AGTGTATTAA	CCAGAACTCT	1200
	GCTGCATCTT	CAATCACTG	GCAGATTCGG	GAGTTAGCCC	GAAGAGTTAA	TACAGAGCAG	1260
	AAACATACCA	CTTTTGAGCA	ACCTGCTCTT	TCAGTTTCAA	AACAGTCACC	ACCAATATCA	1320
	ACATCTAAAT	GTTTTCAGCC	AAAATCTATT	TGTAAGACAC	CAAGCAGCAA	TACCTTGGAT	1380

	GATTACATGA	GCTGTTTATG	AATCCAGTT	GTAAGAAGT	ACTTTCCACC	TGCTTGTCAG	1440
	TTGTCAACAC	CTTATGGCCA	ACCTGCCTGT	TTCCAGCAGC	AACAGCATCA	AATACTTGCC	1500
	ACTCCACTTC	AAAATTTCAT	GTTTATAGCA	TCTTCTTCAG	CAAATGAATG	CATTTGGTT	1560
5	AAAGGAAGAA	TTTATTCAT	TTTAAAGCAG	ATAGGAAGTG	GAGGTTCAAG	CAAGGTATTT	1620
	CAGGTGTTAA	ATGAAAAGAA	ACAGATATAT	GCTATAAAAT	ATGTGAACIT	AGAAGAAGCA	1680
	GATAACCAAA	CTCTGATAG	TTACCGGAAC	GAAATAGCTT	ATTTGAATAA	ACTACAACAA	1740
	CACAGTGATA	AGATCATCCG	ACTTTATGAT	TATGAAATCA	CGGACCACTA	CATCTACATG	1800
	GTAATGGAGT	GTGGAAATAT	TGATCTTAAT	AGTTGGCTTA	AAAAGAAAAA	ATCCATTGAT	1860
10	CCATGGGAAC	GCAAGAGTTA	CTGGAAAAAT	ATGTTAGAGG	CAGTTCACAC	AATCCATCAA	1920
	CATGGCATTG	TTCAAGTGA	TCTTAAACCA	GCTAACTTTC	TGATAGTTGA	TGGAATGCTA	1980
	AAGCTAATTG	ATTTTGGGAT	TGCAAAACCA	ATGCAACAG	ATACAACAAG	TGTTGTTAAA	2040
	GATTCTCAGG	TTGGCAGAGT	TAATTATATG	CCACAGCAAG	CAATCAAAGA	TATGTCTTCC	2100
	TCCAGAGAGA	ATGGGAATAT	TAAGTCAAAG	ATAAGCCCCA	AAAGTGATGT	TTGGTCCTTA	2160
	GGATGTATT	TGTACTATAT	GACTTACGGG	AAAAACCAT	TTCAGCAGAT	AATTAATCAG	2220
15	ATTCTTAAT	TACATGCCAT	AATTGATCCT	AATCATGAAA	TTGAATTTCC	CGATATTCCA	2280
	GAGAAAGATC	TTCAAGATGT	GTTAAAGTGT	TGTTTAAAAA	GGGACCCAAA	ACAGAGGATA	2340
	TCCATTCCG	AGCTCCTGGC	TCATCCCTAT	GTTCAAATTC	AACTCATCC	AGTTAACCAA	2400
	ATGGCCCAAG	GAACCACTGA	AGAAATGAAA	TATGTTCTGG	GCCAACTTGT	TGGTCTGAAT	2460
20	TCTCTAAT	CCATTTGAAA	AGCTGCTAAA	ACTTTATATG	AACACTATAG	TGGTGGTGAA	2520
	AGTCATAAT	CTTCATCCTC	CAAGACTTTT	GAAGAAAAAA	GGGGAAAAAA	ATGA	

Seq ID NO: 647 Protein sequence
Protein Accession #: NP_003309.1

25	1	11	21	31	41	51	
	MESEDLSSRE	LTIDSMNKV	RDINKKFKNS	DLTDELSLNK	ISADTTDMSG	TVNQIMMMAN	60
	NPEDWLSLL	KLEKNSVPLS	DALLNKLIGR	YSQAIELPP	DKYQNESFA	RIQVRFAELK	120
	AIQEPDDARD	YFQMARANCK	KFAFVHISPA	QFELSQGNVK	KSKQLLQKAV	ERGAVPLEML	180
30	EIALRNLMQ	KKQLLSEEEK	KNLSASTVLT	AQESFSGSLG	HLQNRNNSCD	SRGQTTKARF	240
	LYGENMPQD	ASIGYRNSLR	QTNKTKQSCP	FORVPVNLIN	SPDCDVKTDD	SVVPCFMKRG	300
	TSRSECRDLV	VPGSKPSGND	SCELRNLKSV	QNSHFKEPLV	SDEKSSSELI	TDSTLTWNKT	360
	ESSLLAKLEE	TKYEQPEVP	ESNQKQWQSK	RKSECINQNP	AASSNHMQIP	ELARKVNTQ	420
	KHTTFEQPVF	SVSKQSPPI	TSKNFDPKSI	CKTPSSNTLD	DYMSCFRTPV	VKNDFPPACQ	480
35	LSTPYGQPA	FQQQSQPILA	TPLQNLQVLA	SSSANECISV	KGRISILKQ	IGSGSSSKVF	540
	QVLNEKQIY	AIKYVNLLEA	DNQTLDSYRN	BIAYLNKLOQ	HSDKIIRLYD	YEITDQYIYM	600
	VMECGNIDL	SWLKKKSID	PWERKSYWKN	MLEAVHTIQ	HGIVHSDLPK	ANFLIVDGM	660
	KLIDFGIANQ	MQPDTSVVK	DSQVGTVNYM	PPEAIKDMSS	SRENGKSKSK	ISPKSDVWSL	720
40	GCILYMTYG	KTPFQOIINQ	ISKLHAIIDP	NHEIEFPDIP	EKDLQDVLC	CLKRDPKQRI	780
	SIPELLAPHY	VQIQTHPVNQ	MAKGTTEEMK	YVLQQLVGLN	SPNSILKAAK	TLYEHYSGGE	840
	SHNSSSSKTF	EKKRGKK					

Seq ID NO: 648 DNA sequence
Nucleic Acid Accession #: NM_015507
Coding sequence: 241..1902

45	1	11	21	31	41	51	
	CCGCAGAGGA	GCCTCGGCCA	GGCTAGCCAG	GGCGCCCCCA	GCCCTCCCCC	AGGCCGCGAG	60
50	CGCCCCGTCC	CGCGTGCCTC	GCCTCCCTCC	CCAGACTGCA	GGGACAGCAC	CCGGTAACTG	120
	CGAGTGGAGC	GGAGGACCCG	AGCGGCTGAG	GAGAGAGGAG	CGCGCGGCTT	AGCTGCTACG	180
	GGGTCCGGCC	GGCCCTCTCC	CGAGGGGGGC	TCAGGAGGAG	GAAGGAGGAC	CCGTGCGAGA	240
	ATGCCTCTGC	CCTGGAGCCT	TGCGCTCCCG	CTGCTGCTCT	CCTGGGTGGC	AGGTGGTTTC	300
55	GGGAACGCGG	CCAGTGCAAG	GCAATCACGG	TGTTAGCAT	CGGCACGTCA	GCCTGGGGTC	360
	TGTCACTATG	GAATCAAACT	GGCCTGCTGC	TACGGCTGGA	GAAGAAACAG	CAAGGGAGTC	420
	TGTGAAGCTA	CTGCGAACC	TGGATGTAAG	TTTGGTGAGT	CGGTGGGACC	AAACAAATGC	480
	AGATGCTTTC	CAGGATACAC	CGGGAAAAAC	TGCAGTCAAG	ATGTGAATGA	GTGTGGAATG	540
	AAACCCCGGC	CATGCCAACA	CAGATGTGTG	AATACACACG	GAAGCTACAA	GTGCTTTTGC	600
60	CTCAGTGGCC	ACATGCTCAT	GCCAGATGCT	ACGTGTGTGA	ACTCTAGGAC	ATGTGCCATG	660
	ATAAACTGTC	AGTACAGCTG	TGAAGACACA	GAAGAAGGGC	CACAGTGCCT	GTGTCCATCC	720
	TCAGGACTCC	GCCTGGCCCC	AAATGGAAGA	GACTGTCTAG	ATATTGATGA	ATGTGCTCTC	780
	GGTAAAGTCA	TCTGTCCCTA	CAATCGAAGA	TGTGTGAACA	CATTGGAAG	CTACTACTGC	840
	AAATGTCA	TTGGTTTCGA	ACTGCAATAT	ATCAGTGGAC	GATATGACTG	TATAGATATA	900
65	AATGAATGTA	CTATGGATAG	CCATACGTGC	AGCCACCATG	CCAATTGCTT	CAATACCCAA	960
	GGGTCTTCA	AGTGTAAATG	CAAGCAGGGA	TATAAAGGCA	ATGGACTTCG	GTGTTCTGCT	1020
	ATCCCTGAAA	ATTCTGTGAA	GGAAGTCCCT	AGAGCACCTG	GTACCATCAA	AGACAGAATC	1080
	AAGAAGTTGC	TGCTCACA	AAACAGCATG	AAAAAGAAAG	CAAAATTA	AAATGTTACC	1140
	CCAGAACCCA	CCAGACTCC	TACCCCTAAG	GTGAACCTGC	AGCCCTTCAA	CTATGAAGAG	1200
70	ATAGTTTCCA	GAGGCGGAA	CTCTCATGGA	GGTAAAAAAG	GGAAATGAAG	GAATGATA	1260
	GAGGGGCTTG	AGGATGAGAA	AAGAGAAGAG	AAAGCCCTGA	AGAATGACAT	AGAGGAGCGA	1320
	AGCGTCGAG	GAGATGTGTT	TTTCCCTAAG	GTGAATGAAG	CAGTGAAAT	CGGCTTGATT	1380
	CTGCTCCAAA	GGAAAGCGCT	AACTTCCAAA	CTGGAACATA	AAGATTAA	TATCTCGGTT	1440
	GACTGCAGTG	TCAATCATGG	GATCTGTGAC	TGGAAACAGG	ATAGAGAAGA	TGATTTTGAC	1500
75	TGGAATCCTG	CTGATCGAGA	TAATGCTATT	GGCTTCTATA	TGGCAGTTCC	GGCCTTGGCA	1560
	GGTCACAAGA	AAGACATTGG	CCGATTGAAA	CTTCTCTAC	CTGACCTGCA	ACCCCAAGC	1620
	AACCTCTGTT	TGCTCTTTGA	TTACCGGCTG	GCCGGAGACA	AAGTCGGGAA	ACTTCGAGTG	1680
	TTTGTGAAAA	ACAGTAACAA	TGCCCTGGCA	TGGGAGAAGA	CCACGAGTGA	GGATGAAAAG	1740
	TGGAAGACAG	GGAAATTTCA	GTTGTATCAA	GGAACGTGAT	CTACCAAAAG	CATCATTTTT	1800
80	GAAGCAGAAC	GTGGCAAGGG	CAAAACCGGG	GAAATCGCAG	TGGATGGCGT	CTTGCTTGTT	1860
	TCAGGCTTAT	GTCCAGATAG	CCCTTTATCT	GTGGATGACT	GAATGTTACT	ATCTTTATAT	1920
	TTGACTTTGT	ATGTCAGTTC	CCTGGTTTTT	TTGATATTGC	ATCATAGGAC	CTCTGGCATC	1980
	TTAGAATTAC	TAGCTGAAAA	ATTGTAATGT	ACCAACAGAA	ATATTATTGT	AAGATGCCTT	2040
	TCTTGTATAA	GATATGCCAA	TATTGTCTTT	AAATATCATA	TCACTGTATC	TTCTCAGTCA	2100
85	TTTCTGAATC	TTTCCACATT	ATATTATAAA	ATATGGAAAT	GTCACTTTAT	CTCCCTCTCT	2160
	CAGTATATCT	GATTGTGATA	AGTAAGTTGA	TGAGCTTCTC	TCTACAACAT	TTCTAGAAAA	2220
	TATGACATCA	AAGATAGACT	TTTGCCTAAG	TGGCTTAGCT	GGGTCTTTCA	TAGCCAAACT	2280
							2340

WO 02/086443

PCT/US02/12476

TGATATTTA AATCTTTGT AATAATAATA TCCAAATCAT CAAAAA AAAA

Seq ID NO: 649 Protein sequence
Protein Accession #: NP_056322

1	11	21	31	41	51	
MPLFWSLALP	LLLSWVAGGF	GNAASARHHG	LLASARQPGV	CHYGTKLACC	YGWRRNSKGV	60
CEATCEPGCK	FGECVGNKNC	RCPPGYTGKT	CSQDVNECGM	KPRPCQHRVC	NTHGSYKFC	120
LSGHEMLMPDA	TCVNSRTCAM	INCQYSCEDT	EEGPOCLCPS	SGLRLAPNGR	DCLDIDECAS	180
GKVICPYNRR	CVNTFGSYYC	KCHIGFELQY	ISGRYDCIDI	NECTMDSHTC	SHHANCFTNQ	240
GSFKCKCKQG	YKGNGLRCSA	IPENSVKEVL	RAPGTIKDRI	KKLLAHKNSM	KKKAKIKNVT	300
PEPTRTPTPK	VNLQPFNYEE	IVSRGGNSHG	GKKGNEEKMK	EGLEDEKREE	KALKNDIEER	360
SLRGDVFFPK	VNEAGEFGLI	LVQRKALTSK	LEHKDLNISV	DCSFNHGICD	WKQDREDDFD	420
WNPADRDNAI	GFYMAVPALA	GHKKDIGRLK	LLLPDLQPOQ	NFCLLFDYRL	AGDKVKGKLRV	480
PVKNSNNALA	WERTTSEDER	WKTGKIQLYQ	GTDAKTSIIF	EAERGKGTG	EIAVDGVLLV	540
SGLCPDSLIS	VDD					

Seq ID NO: 650 DNA sequence
Nucleic Acid Accession #: NM_003506.1
Coding sequence: 259..2379

1	11	21	31	41	51	
GCAGCTCCAG	TCCCGGACGC	AACCCCGGAG	CCGTCTCAGG	TCCCTGGGGG	GAACGGTGGG	60
TTAGAGCGGG	ACGGGAAGGG	ACAGCGGCCT	TGACCGCCCC	CCCAGTAAT	TGACCCAGGA	120
CTCATTTTCN	GGAAAGCCTG	AAAATGAGTA	AAATAGTGAA	ATGAGGAATT	TGAACATTTT	180
ATCTTTGGAT	GGGGATCTTC	TGAGGATGCA	AAGAGTGATT	CATCCAAGCC	ATGTGGTAAA	240
ATCAGGAATT	TGAAGAAAT	GGAGATGTTT	ACATTTTGT	TGACGTGTAT	TTTCTACCC	300
CTCCTAAGAG	GGCAGCTCT	CTTCACCTGT	GAACCAATTA	CTGTCCAG	ATGTATGAAA	360
ATGGCCTACA	ACATGACGTT	TTTCCTTAAT	CTGATGGGTC	ATTATGACCA	GAGTATTGCC	420
GCGGTGGAAA	TGAGGACATT	TCTTCCTCTC	GCAAACTCGG	AATGTTACCC	AAACATTGAA	480
ACTTCTCTCT	GCAAGCAATT	TGTACCAACC	TGCATAGAAC	AAATTCATGT	GGTTCACCT	540
TGTGTAAC	TTGTGTGAGAA	AGTATATTCT	GATTGCAAAA	AATTAATGA	CACTTTGGG	600
ATCCGATGGC	CTGAGGAGCT	TGAATGTGAC	AGATTACAAT	ACTGTGATGA	GACTGTTCCT	660
GTAACCTTTG	ATCCACACAC	AGAATTTCTT	GGTCTCAGA	AGAAAACAGA	ACAAGTCCAA	720
AGAGACATTG	GATTTTGGTG	TCCAAGGCAT	CTTAAGACTT	CTGGGGGACA	AGGATATAAG	780
TTTCTGGGAA	TTGACCAAGT	TGCGCCTCCA	TGCCCAACAA	TGTATTTTAA	AAGTGATGAG	840
CTAGAGTTTG	CTAGTGTGAT	TATTGGAACA	GTTTCAATAT	TTTGTCTTTG	TGCAACTCTG	900
TTACATTTCC	TTACTTTTTT	AATGTATGTT	AGAAAGATCA	GATACCCAGA	GAGACCAATT	960
ATATATTACT	CTGTCTGTGA	CAGCATTTGA	TCTCTTATGT	ACTTCATTGG	ATTTTGTCTG	1020
GGCGATAGCA	CAGCCTGCAC	TAAGGCAGAT	GAGAAGCTAG	AACCTGGTGA	CAGTGTGTGC	1080
CTAGGCTCTC	AAAATAGGC	TTGACCGGTT	TTGTTCATGC	TTTTGTATTT	TTTCAACAATG	1140
GCTGGCACTG	TGTGTGGGTT	GATCTTACCC	ATTACTTGGT	TCTTAGCTGC	AGGAAGAAAA	1200
TGGAGTTGTG	AAGCCATCGA	GCAAAAAGCA	GTGTGGTTTC	ATGCTGTTGC	ATGGGGAACA	1260
CCAGGTTTCC	TGACTGTTAT	GCTTCTTGCT	CTGAACAAAG	TTGAAGGAGA	CAACATTAGT	1320
GGAGTTTGCT	TTGTTTGGCT	TTATGACCTG	GATGCTTCTC	GCTACTTTGT	ACTCTTGCCA	1380
CTGTGCTTTT	GTGTGTTTGT	TGGGCTCTCT	CTTCTTTTAS	CTGGCATTAT	TTCTCTAAAT	1440
CATGTCGACG	CAAAAGTTT	ACATGATGGC	CGGAACCAAG	AAAACTATAA	GAAATTTATG	1500
ATTCGAATTG	GAGTCTTCAG	CGGCTTGAT	CTTGTGCCAT	TAGTGACACT	TCTCGGATGT	1560
TACGTCTATG	AGCAAGTGAA	CAGGATTACC	TGGGAGATAA	CTTGGGTCTC	TGATCATTTG	1620
CGTCAGTACC	ATATCCCATG	TCCTTATCAG	GCAAAAGCAA	AAGCTCGACC	AGAATTGGCT	1680
TTATTTATGA	TAAATACCTT	GATGACATTA	ATTGTTGGCA	TCTCTGCTGT	CTTCTGGGTT	1740
GGAAAGCAAA	AGACATGCAC	AGAAATGGGCT	GGGTTTTTTA	AACGAAATCG	CAAGAGAGAT	1800
CCAATCAGTG	AAAGTCCGAG	AGTACTACAG	GAATCATGTG	AGTTTTCTTT	AAAGCACAAT	1860
TCTAAAGTTA	AACACAAAAA	GAAGCACTAT	AAACCAAGTT	CACACAAGCT	GAAGTCAATT	1920
TCCAATATCA	TGGGAACCCG	CACAGGAGCT	ACAGCAAAATC	ATGGCACTTC	TGCAGTAGCA	1980
ATTACTAGCC	ATGATTACCT	AGGACAAGAA	ACTTTGACAG	AAATCCAAAC	CTCACAGGAA	2040
ACATCAATGA	GAGAGGTGAA	AGCGGACGGA	GCTAGCACCC	CCAGGTAAAG	AGAACAGGAG	2100
TGTGGTGAAC	CTGCTCTGCC	AGCAGCATCC	ATCTCCAGAC	TCTCTGGGGA	ACAGGTGCGAC	2160
GGGAAGGGCC	AGGCAGGCGAG	TGTATCTGAA	AGTGGCGGGA	GTGAAGGAAG	GATTAGTCCA	2220
AAGAGTGATA	TTACTGACAC	TGGCCTGGCA	CAGAGCAACA	ATTTCAGAGT	CCCCAGTTCT	2280
TCAGAACCAA	GCAGCCTCAA	AGGTTCCACA	TCTCTGCTTG	TTCAACCCAGT	TTCAGGAGTG	2340
AGAAAAGAGC	AGGGAGGTGG	TTGTCAATCA	GATACTTGAA	GAACATTTTC	TCTCGTTACT	2400
CAGAAGCAAA	TTTGTGTTAC	ACTGGAAGTG	ACCTATGCAC	TGTTTTGTAA	GAATCACTGT	2460
TACGTTCTTC	TTTTGCACTT	AAAGTTGCAT	TGCCTACTGT	TATACTGGA	AAAATAGAGT	2520
TCAGAAATAA	TATGACTCAT	TTCAACAAAA	GGTTAATGAC	AACAATATAC	CTGAAAAACG	2580
AAATGTGCAG	GTTAATAATA	TTTTTTTAAAT	AGTGTGGGAG	GACAGAGTTA	GAGGAATCTT	2640
CCTTTTCTAT	TTATGAAGAT	TCTACTCTTG	GTAAGAGTAT	TTTAAGATGT	ACTATGCTAT	2700
TTTACCTTTT	TGATATAAAA	TCAAGATATT	TCTTTGCTGA	AGTATTAAAA	TCTTATCCTT	2760
GTATCTTTTT	ATACATATTT	GAAAAAAGC	TTATATGTAT	TTGAACCTTT	TTGAAATCCT	2820
ATTCAAGTAT	TTTTATCATG	CTATTGTGAT	ATTTTAGCAC	TTTGGTAGCT	TTTACACTGA	2880
ATTCTAAGA	AAATGTGAAA	ATAGTCTTCT	TTTATACTGT	AAAAAAGAT	ATACCAAAAA	2940
GTCTTATAAT	AGGAATTTAA	CTTTAAAAAC	CCACTTATTG	ATACCTTACC	ATCTAAAAATG	3000
TGTGATTTTT	ATAGTCTCGT	TTTAGGAATT	TCACAGATCT	AAATTATGTA	ACTGAAATAA	3060
GGTGCTTACT	CAAGAGGTGT	CCACTATTGA	TGTATTATG	CTGCTCACTG	ATCCTTCTGC	3120
ATATTAAAAA	TAAATGTGCC	TAAAGGGTTA	GTAGACAAAA	TGTAGTCTT	TGTATATTA	3180
GGCCAGTGCC	AATTGACTTC	CCTTTTTTAA	TGTTTCATGA	CCACCCATTG	ATTGTATTAT	3240
AACCACTTAC	ATTGTCTTAT	ATTTTTTGT	TTAATTTTG	TTTCTTAACA	TTTAGAATAT	3300
TACATTTTGT	ATTATACAGT	ACCTTTCTCA	GACATTTTGT	AG		

Seq ID NO: 651 Protein sequence
Protein Accession #: NP_003497.1

1	11	21	31	41	51

WO 02/086443

PCT/US02/12476

MEMFTLLTLC IFPLPLRGHS LFTCEPITVP RCMKAYNMT FFPNLMGHYD QSIAAEMEH 60
 FLPLANLECS PNIEFTFLCKA FVPTCIEQIH VVPPCRKLCE KVYSDCKLI DTFGIRNPBE 120
 LSCDRLOQYCD ETWVVTFFDHP TEFLGPQKKT EQVQRDIGPW CPRHLKTSKG QGYKFLGIDQ 180
 CAPPNPMYF KSELEFAKS FIGTVSIFCL CATLEFLTP LIDVRRFRYP ERPIIYSVC 240
 YSIVSLMYFI GFLGLDSTAC NKADEKLELG DTVVLGSSQNK ACTVLFMLLY FFTMAGTVWN 300
 VILTLTWFLA AGRKWSCEAI BQKAVWFHAV ANGTGPGFLT VMLLANKVEG DNISGVCFVG 360
 LYDLASRYF VLLPLCLCVF VGLSLLLAGI ISLNVHRQVI QHDGRNQEKI KKFMRIGVF 420
 SGLYLVPLVT LLGCYVYEQV NRITWEITWV SDHCROYHIP CPYQAKAKAR PELALFMKY 480
 LMTLIVGISA VFWVGSKKTC TEWAGFFKRN RKRDPISER RVLQESCEPF LKHNSKVHKH 540
 KKHYPSSHK LKVISKSMGT STGATANHGT SAVAITSHDY LGQSTLTIQ TSPETSMREV 600
 KADGASTPRL REQDCGPAS PAASISRLSG EQVDGKGQAG SVSESARSEG RISPKSDITD 660
 TGLAQSNLQ VPSSSEPPSL KGSTSLLVHP VSGVRKEQGG GCHSDT

Seq ID NO: 652 DNA sequence
 Nucleic Acid Accession #: NM_014791.1
 Coding sequence: 171..2126

1 11 21 31 41 51
 20 TTGGCGGGCG GAAGCGGCCA CAACCCGGCG ATCGAAAAGA TTCTTAGGAA CGCCGTACCA 60
 GCCCGCTCTC TCAGACACAG AGGCCCTCTG CCTTCTCTCG GCGCGCGCTC AGCCGTGCCC 120
 TCGCCCTCTC AGGTTCCTTT TCTAATTCCA AATAAACTTG CAAGAGGACT ATGAAAGATT 180
 ATGATGAAC TCTCAAATAT TATGAATTAC ATGAACTAT TGAGACAGGT GCGCTTGCAA 240
 AGGTCAAAC TGCTGCCAT ATCCTTACTG GAGAGATGGT AGCTATAAAA ATCATGGATA 300
 25 AAAACACACT AGGGAGTGAT TTGCCCCGGA TCAAAAACGGA GATTGAGGCC TTGAAGAACC 360
 TGAGACATCA GCATATATGT CAACTCTACC ATGTGCTAGA GACAGCCAAC AAAATATTCA 420
 TGGTTCTTGA GTACTGCCCT GGAGGAGAGC TGTGTGACTA TATAATTTCC CAGGATCGCC 480
 TGTCAGAAGA GGAGACCCGG GTTGTCTTCC GTCAGATAGT ATCTGCTGTT GCTTATGTGC 540
 ACAGCCAGGG CTATGCTCAC AGGGACCTCA AGCCAGAAAA TTTGCTGTTT GATGAATATC 600
 30 ATAAATTAAA GCTGATTGAC TTTGGTCTCT GTGCAAAACC CAAGGGTAAC AAGGATTACC 660
 ATCTACAGAC ATGCTGTGGG AGTCTGGCTT ATGCAGCACC TGAGTTAATA CAAGGCAAT 720
 CATATCTTGG ATCAGAGGCA GATGTTTGGG GCATGGGCAT ACTGTTATAT GTTCTTATGT 780
 GTGGATTCT ACCATTGAT GATGATAATG TAATGGCTTT ATACAAGAG ATTATGAGAG 840
 35 GAAAAATATG TGTTCCTCAAG TGGCTCTCTC CCAGTAGCAT TCTGCTCTT CAACAATGC 900
 TGCAGGTGGA CCCAAGAGAA CCGATTCTTA TGAAAAATCT ATTGAACCAT CCCTGGATCA 960
 TGCAGATTGA CAATCTCTCT GTTGAGTGGC AAAGCAAGAA TCCTTTTATT CACCTCGATG 1020
 ATGATTGCGT AACAGAACTT TCTGTACATC ACAGAAACAA CAGGCAACAA ATGGAGGATT 1080
 TAATTTCACT GTGCGAGTAT GATCACTCA CCGCTACCTA TCTTCTGCTT CTAGCCAGA 1140
 40 AGGCTCGGGG AAAACCACTT CGTTTAAGGC TTTCTTCTTT CTCTGTGGA CAAGCCAGTG 1200
 CTACCCCAT CACAGACATC AAGTCAAATA ATTGGAGTCT GGAAGATGTG ACCGCAAGTG 1260
 ATAAAAATTA TGTGGCGGGA TTAATAGACT ATGATTGGTG TGAAGATGAT TTATCAACAG 1320
 GTGCTCTGAC TCCCGAACA TCACAGTTTA CCAAGTACTG GACAGAATCA AATGGGTGG 1380
 AATCTAAATC ATTAACCTCA GCCTTATGCA GAACACCTGC AAATAAATTA AAGAACAAAG 1440
 AAAATGTATA TACTCTTAAG TCTGCTGTAA AGAATGAAGA GTACTTTATG TTTCTGAGC 1500
 45 CAAAGACTCC AGTTAATAAG AACCAGCATA AGAGAGAAAT ACTCACTACG CCAATCTGTT 1560
 ACACCTACCC CTCAAAAGCT AGAAACCACT GCGTGAAAGA AACTCCAATT AAAATACAG 1620
 TAAATCTAGT AGGAACAGAC AAGTTAATGA CAGGTGTCTT TAGCCCTGAG AGGCGGTGCC 1680
 GCTCAGTGA ATTGATCTCT AACCAAGCAC ATATGGAGGA GACTCCAAAA AGAAGGGAG 1740
 50 CCAAGATGTT TGGGAGCCTT GAAAGGGGGT TGGATAAGGT TATCACTGTG CTACCCAGGA 1800
 GCAAAAGGAA GGGTTCTGCC AGAGACGGGC CCAGAAGACT AAAGCTTCAC TATAATGTGA 1860
 CTACCACTAG ATTAGTGAAT CCAGATCAAC TGTGTAATGA AATAATGTCT ATCTCTCAA 1920
 AGAAGCATGT TGACTTTGTA CAAAAGGGTT ATCACTGAA GTGTCAAACA CAGTCAGATT 1980
 TTGGGAAAGT GACAATGCAA TTTGAATTAG AAGTGTGCCA GCTTCAAACA CCGCATGTGG 2040
 TGGGTATCAG GAGGACGCGG CTTAAGGGCG ATGCTGTGGT TTACAAAAGA TTAGTGAAG 2100
 55 ACATCCTATC TAGCTGCAAG GTATAATTGA TGGATTCTTC CATCTGCCG GATGAGTGTG 2160
 GGTGTATAC AGCTCAATGA AAGACTGTTA TGATCGCTTT GATTTTAAAG TTCATTGGAA 2220
 CTACCAACTT GTTCTTAAAG AGCTATCTTA AGACCAATAT CTCCTTGTGT TAAACAAAA 2280
 GATATTATTT TGTGTATGAA TCTAAATCAA GCGCATCTGT CATTATGTTA CTGCTTTTT 2340
 60 TAATCATGTG GTTTGTGATA TTAATAATTG TTGACTTTCT TAGATTCACT TCCATATGTG 2400
 AATGTAAGCT CTTAACATAG TCTCTTTGTA ATGTGTAATT TCTTCTGAA ATAAACCAT 2460
 TTGTGAATAT

Seq ID NO: 653 Protein sequence
 Protein Accession #: NP_055606.1

1 11 21 31 41 51
 65 MKDYDELLKY YELHETIGTG GFAKVKLACH ILTGEMVAIK IMDKNTLGS DLPRIKTEIEA 60
 LKNLRHQHIC QLYHVLETAN KIPMVLEYCP GGELPDYIIS QDRLESEETR VVFRQIVSAV 120
 70 AYVHSQGYAH RDLKPENLLF DEYHKLKLID FGLCAKPKGN KDYHLQTCGG SLAYAAPELI 180
 QGKSYLGSEA DVWSMGLILY VLMCGFLPFD DDNVMAIYKK IMRGKYDVPK WLPSSSILL 240
 QQMLQVDPKK RISMKNLNLH PHIMQDYNYP VEWQSKNFFI HLDDDCVTEL SVHHRNRQT 300
 MEDLISLWQY DHLTATYLLL LAKKARGKPV RLRLSSFSFG QASATPFTDI KSNNSLEDV 360
 75 TADKHYVAG LIDYDWCEDD LSTGAATPRT SOFTKYWTES NGVESKSLTP ALCRTPANKL 420
 RNKENVYTK SAVKNEEYPM FPEKTPVKN NQHKREILTT PNRYTTPSKA RNQCLKETPI 480
 KIPVNSTGTD KLMTGVISPE RRCRSVELDL NQAHMEETPK RKGAKVPGSL ERGLDKVITV 540
 LTRSKRKGSA RDGPRRLKLH YNVTTTLVN PDQLLNEIMS ILPKKHVPV QGYTLKQCT 600
 QSDFGKVTMQ FELEVCLQK PDVVGIRRRQ LKGDWVYKR LVEDILSSCK V

Seq ID NO: 654 DNA sequence
 Nucleic Acid Accession #: NM_000582
 Coding sequence: 88..990

1 11 21 31 41 51
 85 GCAGAGCACA GCATCGTCGG GACCAGACTC GTCTCAGGCC AGTTGCAGCC TTCTCAGCCA 60
 AACGCCGACC AAGGAAAACT CACTACCATG AGAATTGCAG TGATTGCTT TTGCTCCTCA 120

WO 02/086443

5
10
15
20
25

```

GGCATCACT GTGCCATACC AGTTAAACAG GCTGATTCTG GAAGTCTCTGA GGAAAAGCAG 180
CTTTACAACA AATACCCAGA TGCTGTGGCC ACATGGCTAA ACCCTGACCC ATCTCAGAAG 240
CAGAATCTCC TAGCCCCACA GACCCCTCCA AGTAAGTCCA ACCAAAGCCA TGACCACATG 300
GATGATATGG ATGATGAAGA TGATGATGAC CATGTGGACA GCCAGGACTC CATTGACTCG 360
AACGACTCTG ATGATGTAGA TGACACTGAT GATTCTCACC AGTCTGATGA GTCTCACCAT 420
TCTGATGAAT CTGATGAAGT GGTCACTGAT TTTCCCAACG ACCTGCCAGC AACCGAAGTT 480
TTCACTCCAG TTGTCCCCAC AGTAGACACA TATGATGGCC GAGGTGATAG TGTGTTTAT 540
GGACTGAGGT CAAAATCTAA GAAGTTTCGC AGACCTGACA TCCAGTACCC TGATGCTACA 600
GAGGAGGACA TCACCTCACA CATGGAAGC GAGGAGTTGA ATGGTGCATA CAAGGCCATC 660
CCCGTTGCC AGGACCTGAA CGCGCCTTCT GATTGGGACA GCCGTGGGAA GGACAGTTAT 720
GAAACGAGTC AGCTGGATGA CCAGAGTGCT GAAACCCACA GCCACAAGCA GTCCAGATTA 780
TATAAGCGGA AAGCCAATGA TGAGAGCAAT GAGCATTCCG ATGTGATTGA TAGTCAGGAA 840
CTTTCCAAAG TCAGCCGTGA ATTCCACAGC CATGAATTC ACAGCCATGA AGATATGCTG 900
GTTGTAGACC CCAAAGTAA GGAAGAAGAT AAACACCTGA AATTTCGTAT TTCTCATGAA 960
TTAGATAGTG CATCTCTGA GGTCAATTAA AAGGAGAAAA AATACAATT CTCACCTTGC 1020
ATTAGTCAAA AAGAAAAAT GCTTTATAGC AAAATGAAAG AGAACATGAA ATGCTTCTTT 1080
CTCAGTTTAT TGGTTGAATG TGATCTATT TGAGTCTGGA AATAACTAAT GTGTTTGATA 1140
ATTAGTTTAG TTTGTGGCTT CATGGAACCT CCCTGTAAC TAAAGCTTC AGGGTTATGT 1200
CTATGTTTAT TCTATAGAG AATGCAAACT TATCACTGTA TTTTAATATT TGTATTCTC 1260
TCATGAATAG AATTTTATGT AGAAGCAAACT AAAATACTTT TACCCACTTA AAAGAGAAAT 1320
ATAACATTTT ATGCTACTAT AATCTTTTGT TTTTAAAGTT AGTGTATATT TGTGTTGTG 1380
TATCTTTTGT TGGTGTGAAT AATCTTTTGA TCTGAATGT AATAAGAATT TGGTGTGTG 1440
AATGTCTTAT TTGTTTCCG ACGGTGTGCC AGCAATTAAT AAAACATAAC CTTTTTACT 1500
GCCTAAAAA AAAAAAAAAA AAAA

```

Seq ID NO: 655 Protein sequence
Protein Accession #: NP_000573

30
35

```

1 11 21 31 41 51
MRIAIVICFL LGITCAIPVK QADSGSSEK QLYNKYPDAV ATWLNPDPSQ KQNLAPQTL 60
PSKSNESHDE MDMDDDDDD DHVDSQDSID SNDSDDVDDT DSHQSDSH HSDEDELVT 120
DFPTDLPAE VFPTVPVTV TYDGRGDSVV YGLRSKSKF RRPDIQYPA TDEDITSHME 180
SEELNGAYKA IPVAQDLNAP SDWDSRGKDS YETSQLDDQS AETHSHKQSR LYKRKANDES 240
NEHSDVIDSQ ELISKVSRFPH SHEFPHSHEDM LVVDPKSKEE DKHLKFRISH ELDSASSEVN

```

Seq ID NO: 656 DNA sequence
Nucleic Acid Accession #: NM_003108.1
Coding sequence: 76..1401

40
45
50
55
60
65
70
75

```

1 11 21 31 41 51
GGGGTGGGAG GGGGAGGGGG ACCTCCGCAC GAGACCCAGC GGCCCGGGTT GGAGCGTCCA 60
GCCCTGCAAC GGATCATGGT GCAGCAGGCG GAGAGCTTGG AAGCGGAGAG CAACCTGCCC 120
CGGGAGCGGC TGGACACGGA GGAGGGGGAA TTATATGGCT GCAGCCCGGT GCCCTGGAC 180
GAGAGCGACC CAGACTGGTG CAAGACGGCG TCGGGCCACA TCAAGCGGCC GATGAACGCG 240
TTCATGGTAT GGTCCAAGAT CGAACGCGAG AAGATCATGG AGCAGTCTCC GGACATGCAC 300
AACGCCGAGA TCTCCAAGAG CCGTGGCAAG CGCTGGAAAA TGCTGAAGGA CAGCGAGAAG 360
ATCCCGTTCA TCGGGGAGGC GGAGCGGCTG CGGCTCAAGC ACATGGCCGA CTACCCCGAC 420
TACAAGTACC GGCCCGCGAA AAAGCCCAAA ATGGACCCCT CGGCCAAGCC CAGCGCCAGC 480
CAGAGCCCAAG AGAAGAGCGC GGCGCGCGGC GCGCGCGGGA GCGCGGGCGG AGGCGCGGGC 540
GGTGCCAAAG CCTCAAAGGG CTCAGCAAG AAATGCGGCA AGCTCAAGGC CCGCGCGGCC 600
GCGGGCGCCA AGSCGGGCGC GGGCAAGGCG GCCCAGTCCG GGGACTACGG GGGCGCGGGC 660
GACGACTACG TGCTGGGCGC CCGCGGCTG AGCGCTCGG GCGGGCGCGG CGCGGGCAAG 720
ACGGTCAAGT CGGTGTTTCT GGATGAGGAC GACGACGAG ACGACGACGA CGACGAGCTG 780
CAGCTGCAGA TCAACACAGG GCCGACGAG GAGGACGAG AACCAACGCA CCAGCAGCTC 840
CTGCAGCGCG CGGGCGCAGC GCCGTGCGAG CTGCTGAGAC GCTACAACGT CGCCAAAGTG 900
CCCGCCAGCC CTAGCTGAG CAGCTCGGCG GAGTCCCGCG AGGGAGGAGG CCTCTACGAC 960
GAGGTGCGGG CGCGCGCGAC CTCGGGCGCC GGGGCGGCGA GCCGCTCTA CTACAGCTTC 1020
AAGAACATCA CCAAGCAGCA CCCGCGCGCG CTCGCGCAGC CCGCGCTGTC GCCCGCTGCC 1080
TCGCGCTCGG TGTCCACCTC CTGCTCCAGC AGCAGCGGCA GCAGCAGCGG CAGCAGCGCG 1140
GAGGACGCGC AGCAGCTGAT GTTCCAGCTG AGCTTGAATT TCTCTCAAAG CGCGCACAGC 1200
GCCAGCGAGC AGCAGCTGGG GGGCGGCGCG GCGGCGGCGA ACCTGTCCCT GTGCTGGTG 1260
GATAAGGATT TGGATTCGTT CAGCGAGGCG AGCCTGGGCT CCCACTTCCA GTTCCCGGAC 1320
TACTGCACGC CGGAGCTGAG CGAGATGATC GCGGGGAGCT GGCTGGAGGC GAACCTTCTC 1380
GACCTGGTGT TCACATATTG AAAGCGCGCC GCTGCTCGCT CTTTCTCTCG GAGGGTGCAG 1440
AGCTGGGTTT CTTGGGAGGA AGTTGTAGTG GTGATGATGA TGATGATGAT AATGATGATG 1500
ATGATGGTGG TGTGTAGTGT GCGGTGGTA GGGTGGAGGG GAGAGAAGAA GATGCTGATG 1560
ATATTGATAA GATGTCGTGA CGCAAAGAAA TTGGAAGAAA TGATGAAAT TTTGGTGGAG 1620
TTAAAGTGAA ATGAGTAGTT TTTAAACATT TTCTCTGTCC TTTTGTGTC CCCCTCCCT 1680
TCCTTTATCG TGTCTCAAGG TAGTTGCATA CCTAGTCTGG AGTTGTGATT ATTTTCCCAA 1740
AAAATGTGTT TTTGTAATTA CTATTCTTT TTCTGAAAT TCGTGATTGC AACAAAGGCA 1800
GAGGGGGCGG CGCGCGCGAG GGGAGGTAGG ACCCGCTCG GAAGGCGCTG TTTGAAGCTT 1860
GTCGGTCTTT GAAGTCTGGA AGACGTCTGC AGAGGACCTT TTTGGCAGCA CAACTGTTC 1920
TCTAGGGAGT TGGTGGAGAT ATTTTTTTTT CTTAAGAGAA CTTAAGAAGC TGGTGATTTT 1980
TTTTTAACA AAAAAAGG

```

80
Seq ID NO: 657 Protein sequence
Protein Accession #: NP_003099.1

85

```

1 11 21 31 41 51
MVQQAESLEA ESNLPREALD TEEGEFMACS FVALDESDPD WCKTASGHIK RPMNAFMVNS 60
KIERRKIMEQ SPDMINAEIS KRLGKRWML KDSEKIPFIR EAERLRLKHM ADYPDYKYRP 120
RKPKMPSA KPSASQSPK SAAGGGGSA GGGAGGAKTS KGSKKCKGL KAPAAAGAKA 180
GAGKAAQSGD YGGAGDDTVL GSLRVSGSGG GGAGKTVKCV FLDEDDDDDD DDELQLQIK 240

```

WO 02/086443

PCT/US02/12476

QEPDEDEEP PHQQLLPQP QPPSLLRRY NVAKVPASPT LSSSAESPEG ASLYDEVRA 300
ATSGAGGSR LYYSFRNITK QHPPPLAQPA LSPASSRSVS TSSSSSSGSS SGSSGEDADD 360
LMFPLSLNFS QSAHSASEQQ LGGGAAAGNL SLSLVDRDLD SPSEGLSGSH FEPPDYCTPE 420
LSEMIAGDNL EANFSDLVFT Y

Seq ID NO: 658 DNA sequence
Nucleic Acid Accession #: NM_001719
Coding sequence: 123..1418

10 1 11 21 31 41 51
GGGCGCAGCG GGGCCCGTCT GCAGCAAGTG ACCGACGGCC GGGACGGCCG CCTGCCCCCT 60
CTGCCACCTG GGGCGGTGCG GGCCCGGAGC CCGGAGCCCG GGTAGCGCGT AGAGCCGGCG 120
CGATGCACGT GCGCTCACTG CGAGCTGCGG CGCCGCACAG CTTCGTGGCG CTCTGGGCAC 180
15 CCTGTTCTCT GCTGGCTCTC GGCCCTGGCG ACTTCAGCCT GGACAAACGAG GTGCACTCGA 240
GCTTCATCCA CCGCGGCTC CGCAGCCAGG AGCGCGGGGA GATGCAGCGC GAGATCCTCT 300
CCATTTTGGG CTTCGCCAC CGCCCGCGCC CGCAGCTCCA GGGCAAGCAC AACTCGGCAC 360
CCATGTTTCT GCTGGACCTG TACAAGCCA TGGCGGTGGA GAGGGGCGG GGGCCCGCGG 420
GCCAGGGCTT CTCTACCCC TACAAGCCG TCTTCAGTAC CCAGGGCCCC CCTCTGGCCA 480
20 GCTTCAAGA TAGCCATTTC CTCACCGACG CGACATCGT CATGAGCTTC GTCAACCTCG 540
TGGAACTGA CAAGGAATTC TTCAACCCAC GCTACCACCA TCGAGAGTTC CGGTTTGATC 600
TTTCCAAGAT CCCAGAAGGG GAAGCTGTCA CGGCAGCCGA ATTCCGGATC TACAAGGACT 660
ACATCCGGGA AGCTTCGAC AATGAGAGCT TCGGATCAG: CGTTTATCAG GTGCTCCAGG 720
AGCACTTGGG CAGGGAATCG GATCTCTTCC TGCTCGACAG CGTACCCTC TGGGCTCGG 780
25 AGGAGGGCTG GCTGTGTTT GACATCACAG CCACCAGCAA CCACTGGGTG GTCAATCCGC 840
GGCAACACT GGGCTCGAG CTCTCGGTGG AGACGTGGA TGGCGAGAGC ATCAACCCCA 900
AGTTGGGGG CTTGATTGGG CGGCACGGGC CCCAGAACAA CGAGCCCTTC ATGGTGGCTT 960
TCTTCAAGGC CACGAGGTC CACTTCCGCA GCATCCGTC CACGGGAGAG AAACAGCGCA 1020
GCCAGAACCG CTCCAAGACG CCCAAGAAC AGGACGGCCT GTAAGAAGCA CGAGCTGTAT GTCAACCTCC 1140
30 AGAACAGCAG CAGCGACCG AGGCAGGCCT GTAAAGAAC CAAGCTGTAT GTCAACCTCC 1200
GAGACCTGGG CTGGCAGGAC TGGATCATCG CGCTGAAGG CTACGCCGCC TACTACTGTG 1260
AGGGGGAGTG TGCCTTCCCT CTGAACCTCT ACATGAACGC CACCAACCAAC GCCATCGTGC 1320
AGACCTGCTT CCACTTCATC AACCOCGAAA CGGTGCCCAA GCCCTGCTGT GCGCCCAAGC 1380
AGCTCAATGC CATCTCCGTC CTCTACTTCG ATGACAGCTC CAACGTCTAT CTGAAGAAAT 1440
35 ACAGAAACAT GGTGTCCGG GCCTGTGGCT GCCACTAGCT CCTCCGAGAA TTCAAGACCT 1500
TTGGGGCCAA GTTTTCTCG ATCTCCATT GCTCGCCTTG GCCAGGAACC AGCAGACCAA 1560
CTGCTTTTG TGAGACCTTC CCTCCCTAT CCGCAACTTT AAAGGTGTGA GAGTATTAGG 1620
AAACATGAGC AGCATATGGC TTTTGATCAG TTTTTCAGTG GCAGCATCCA ATGAACAAGA 1680
TCTCAAGC TGTCAGGCA AAACCTAGCA GGAAAAAACA ACAACGCATA AAGAAAAATG 1740
40 GCGGGCCAG GTCATTGCTC GGAAGTCTC AGCCATGCAC GGAAGTCTT CCAAGGTAA 1800
TTATGAGCGC CTACAGCCA GGCACCCAG CCGTGGGAGG AAGGGGGCGT GGCAAGGGGT 1860
GGCACATTG GTGTCTGTGC GAAAGGAAAA TTGACCCGGA AGTTCTCTGA ATAAATGTCA 1860
CAATAAACG AATGAATG

Seq ID NO: 659 Protein sequence
Protein Accession #: NP_001710

1 11 21 31 41 51
50 MHVRLRAAA PHSFVALWAP LFLRLSALAD FSLDNEVHSS FIHRLRSQE RREMQREILS 60
ILGLPHRRPRL HQGKHNSAP MFMLDLNLYM AVEEGGGPGG QGFSYPYKAV FSTQGPPLAS 120
LQSHFLIDA DMVMGFVNLV EHDKEFFHPR YHREFFRDL SKIPEGAVT AAEFRYKDY 180
IRERDNETF RISVQVLQE HLGRESLFL LDSRTLWASE EGWLVFDITA TSNHVVNPR 240
HNLGLQLSVE TLDQGSINPK LAGLIGRHGP QNKQPFMVAF FKATBVHFRS IRSTGSKQRS 300
55 QNRKTPKQ EALRMANVAE NSSDQRQAC KKHLYVSFR DLGWQDWIIA PEGYAAYCE 360
GECAPPLNSY MNATNHAIVQ TLVHFINPET VPKPCCAPTQ LNAISVLYFD DSSNVILKKY 420
RNMVVRACGC H

Seq ID NO: 660 DNA sequence
Nucleic Acid Accession #: Eps sequence
Coding sequence: 211..1895

65 1 11 21 31 41 51
GGATCTGAGG GGGGCCAGT CACTTCTCC ACGTTCTCGT GCTGGGCGGG AGGAGGGGAT 60
GGGGCTTGGG AGGCAGCCTG CTCTCCAGTC CCTATCCACC CACAGGTTTT TTGGGTGGGA 120
GAGGAATTAT CTGATAAAAT TCCTGGGTGA ATATTTTAA AAACGGAGAG TTTTAAAAA 180
TGATTTTTT CCCTCGAAA TGACCTTTT ATGCTTCGAA CGAGTTGTG AAACAGCATA 240
70 GTGCTTTTC TTTTCTCTC TTTTCTACG ATAAATGAAA GCATTTCTTC AAGAAAAAGG 300
CACAGTTCC TTGAACAGCT GGATTCTGAT GGCACCAITA CTATAGAGGA GCAGTTGTG 360
CTTGCTGTA AAGCGAAAGT ACAATGTGAA CTCAACATCA CAGCTCAACT CCAGGAGGGA 420
GAAGGTAATT GTTCCCTGA ATGGGATGGA CTCAATTGTT GGGCCAGAGG AACAGTGGG 480
AAAATATCG CTGTTCCATG CCTCCTTAT ATTTATGACT TCAACATAA AGGAGTTGCT 540
75 TTCCGACACT GTAACCCCAA TGGAAACATG GATTTATGC ACAGCTTAAA TAAACATGG 600
GCCAATTATT CAGACTGCTC TCGCTTTCTG CAGCCAGATA TCAGCATAGG AAAGCAAGAA 660
TTCTTTGAAC GCCTCTATGT AATGTATACC GTTGGTACT CCATCTCTTT TGGTTCTCTG 720
GCTGTGCTTA TTCTCATCAT TGGTACTTC AGACGATTGC ATTGCACTAG GAATATATC 780
CACATGCACT TATTGTGTC TTTCTGCTG AGAGCTACAA GCATCTTTGT CAAAGACAGA 840
GTAGTCCATG CTCACATAGG AGTAAAGGAG CTGGAGTCCC TAATAATGCA GGATGACCCA 900
80 CAAAATTCCA TTGAGGCAAC TTCTGTGGAC AAATCACAAT ATATCGGGTG CAAGATTGCT 960
GTTGTGATGT TATTTACTT CTTGGCTACA AATTATTATT GGATCTCGT GGAAGGTCTC 1020
TACTGCTATA TTCTCATCTT TGTGGCTTTC TTTTCGAGCA CCAAAATAC GTGGGGCTTC 1080
ATCTTGATAG CTTGGGGGTT TCCAGCAGCA TTTGTTGAG CATGGGCTGT GGCACGAGCA 1140
85 ACTCTGGCTG ATGCGAGGTG CTGGGAACCT AGTGCTGGAG ACATCAAGTG GATTATCAA 1200
GCACCGATCT TAGCAGTAT TTTATTCTGT TTCTGAATAC GGTTAGAGTT 1260
CTAGCTACCA AATCTGGGA GACCAATGCA GTTGGCATG ACACAAGGAA GCAATACAGG 1320
AAACTGGCCA AATCGACACT GGTCTGTGTC CTAGTCTTTG GAGTGCAATTA CATCGTGTTC 1380

WO 02/086443

PCT/US02/12476

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80

Seq ID NO: 661 Protein sequence
 Protein Accession #: E05 sequence

1 11 21 31 41 51
 MLRSSLSSTSI VLFLPSSFSST INESISSRKR HRFLEQLDSD GTITIEBQIV LVLKAKVQCE 60
 LNITAIQLQEG EGNCFPEWDG LICWPRGTVG KISAVPCPPY IYDFNHKGVA FRHCNPNGTW 120
 DFMHSLNKTW ANYSDCLRFL QPDISIGKQE FFERLYVMYV VGYISIFGSL AVAILIIGYF 180
 RRLHCTRNHY RMHLFVSFML RATSIFVKDR VVHAHIGVKE LESLIMQDDP QNSIEATSV 240
 KSYQYICKIA VMFIYFLAT NYWILVEGL VLNHLIFVAF FSDTKYLWGF ILIGWGFPA 300
 FVAAWAVARA TLADARCWEL SAGDIKWIYQ APILAAIGLN FILPLNTVRV LATKIWTNA 360
 VGHDTKQYR KLAKSTLVLV LVFGVHYIVF VCLPHSFTGL GWEIRMHCEL FFNSFQGFV 420
 SIYYCYNGE VQAEVKQWMS RWNLSVDWKR TPFCGSRRCG SVLTVTHST SSQSQVAAS 480
 RMVLISGKAA KIASRQPDH ITLPGYVWSN SEQDCLPHSF HEETKEDSGR QGDDILMEKP 540
 SRPMESNPDT EGCQGETEDV L

Seq ID NO: 662 DNA sequence
 Nucleic Acid Accession #: NM_005048
 Coding sequence: 143..1795

1 11 21 31 41 51
 GGCCTGGTGC CCGGGCCCGA CCACCCAGC TGCGGTCGT TACTGGCCAC AAGTTTGCTC 60
 TGGCCAGCCG AAGTTGGCAA CTTGGAAGCT TCTCCCGGGC TCTGGAGGAG GGTCCCTGCT 120
 TCTTCTTACA GCCCTTCCGG GCATGGCCGG GCTGGGGGGG TCGTCCACG TCTGGGGTTG 180
 GCTAATGCTC GGCAGCTGCC TCCTGGCCAG AGCCCAAGCTG GATTCTGAT GCACCAATAC 240
 TATAGAGGAG CAGATTGTCC TTGTGCTGAA AGCGAAAGTA CAATGTGAAC TCAACATCAC 300
 AGCTCAACTC CAGAGGGGAG AAGGTAATG TTCCCTGAA TGGGATGGAC TCATTGTGTT 360
 GCCCAGAGGA ACAGTGGGGA AATATCGGC TGTTCATGC CCTCCTTATA TTTATGACTT 420
 CAACCATAAA GGAGTTGCTT TCCGACACTG TAACCCCAAT GGAACATGGG ATTTTATGCA 480
 CAGCTTAAAT AAAACATGGG CCAATTATTC AGACTGCCTT CGCTTTCTGC AGCCAGATAT 540
 CAGCATAGGA AAGCAAGAT TCTTTGAACG CCTCTATGTA ATGTATACCG TTGGCTACTC 600
 CATCTCTTTT GGTTCCCTTG CTGTGGCTAT TCTCATCAT GGTACTTCA GACGATTGCA 660
 TTGCATAGG AACTATATCC ACATGCACCT ATTTGTGTCT TCTATGCTGA GAGCTACAAG 720
 CATCTTTGTC AAAGACAGAG TAGTCCATGC TCACATAGGA GTAAAGGAGC TGGAGTCCCT 780
 AATAATGCAG GATGACCCAC AAAATTCCAT TGAGGCAACT TCTGTGACA AATCACAATA 840
 TATCGGGTGC AAGATTGCTT TTGTGATGTT TATTACTTTC CTGGCTACAA ATTATTATTG 900
 GATCCTGGTG GAAGGTCTCT ACCTGCATAA TCTCATCTT GTGGCTTCTT TTTCCGACAC 960
 CAAATACCTG TGGGGCTTCA TCTTGATAGG CTGGGGGTTT CCAGCAGCAT TTGTTGCAAG 1020
 ATGGGCTGTG GCACGAGCAA CTCTGGCTGA TGGCAGGTGC TGGGAACTTA GTGCTGGAGA 1080
 CATCAAGTGG ATTTATCAAG CACCGATCTT AGCAGCTATT GGGCTGAATT TTAATCTGTT 1140
 TCTGAATACG GTTAGAGTTC TAGCTACCAA AATCTGGGAG ACCAATGCAG TTGGGCATGA 1200
 CACAAGGAAG CAATACAGGA AACTGGCCAA ATCGACACTG GTCCCTGGTCC TAGTCTTTGG 1260
 ATGCACTTAC ATCGTGTGCG TATGCCTGCC TCACTCCTTC ACTGGGCTCG GGTGGGAGAT 1320
 CCGCATGCAC TGTGAGCTCT TCTCAACTC CTTTCAGGGT TTTCTTGTGT CTATCATCTA 1380
 CTGCTACTGC AATGGAGAGG TTCAGGCAGA GGTGAAGAAG ATGTGGAGTC GGTGGAATCT 1440
 CTCCGTGGAC TGGAAAAGGA CACCGCCATG TGGCAGCCCG AGATCCGCTC CAGTGTCTAC 1500
 CACCGTGAGC CACAGCACCA GCAGCCAGTC ACAGTGGGCG GCCAGCACAC GCATGGTGTCT 1560
 TATCTCTGCG AAGCTGCGCA AGATCGCCAG CAGACAGCTC GACAGCCACA TCACTTTACC 1620
 TGGCTATGTC TGGAGTAATC CAGAGCAGGA CTGCCTGCCA CACTCTTTCC ACGAGGAGAC 1680
 CAAGGAAGAT AGTGGGAGGC AGGAGATGTA TATTCTAATG GAGAAGCCTT CCAGGCCTAT 1740
 GGAATCTAAC CCAGACACTG AAGGATGCCA AGGAGAAACT GAGGATGTTT TCTGAATGGA 1800
 CATTTGTGGC TGACTTTTCT GGGCTGGTCC AATGGCTGGT TGTGTGAGAG GGCCTGGCTG 1860
 ATACTCCTAT ATCTGAGTTC AAAGGCTGAA AATTACAGTA AGGTGTTACT TAATAATAGT 1920
 TTTTAGGCTC CATGAATTGG CTCCTGTAAA TACTAACGAC ATGAAAATGC AAGTGTCAAT 1980
 GGAGTAGTTT ATTACCTTCT ATTGGCATCA AGTTTTCCTC TAAATTAATG TATGTTATTT 2040
 GCTCTGTGAT TGTTCATTTT TTTCTGTAC TTTTGGGTAG AAAAAGATT CAATGTCTTG 2100
 GCTGTAGCTT TCTCTCATAT ATATCACCTT AAATATAATG AAGATCTTTT AGTGTGTATC 2160
 ATTTTCTTTT TAGAACTAG TATTCTCTTA TTTCTTACTT TAATGTACTT CTATCACTGC 2220
 ATTTATTTTG CCTGTGCATA GGAGCAATTA GGATCTAAAA AAATATATGG GAAGATAAAA 2280
 GATCTAAGAA CAAGTACTTG CTGGAAAAAT AGTTGGCTGG ACATGTATAA AATAATGCAT 2340
 TTATAACAAT TACATGTGTT TTTGGGAACA AGGAAAAATT CTCAAAAAAG AATATTTTCA 2400
 ACATCCCTTC TTTTGAATGG CCTCTTTGTG ACCAGCCAGA CCTCAGGTCT TCACTCTTTC 2460
 TTCTTTGTA ACCATGTCAT GTGGAAAGAT TTCCTCAGTT AGTGAGCTTG TGTCTGCAAA 2520
 TTGATTTTGT TTGTAATGTA TTTGATAGC AAATCATGCT GCATCTATAT CTTTCTCTG 2580
 TTTGAGCTGT TACTACATTG TACATGGCAT GTGGGATCAA TAAAAAATTT GTTTTAAAAA 2640
 T

Seq ID NO: 663 Protein sequence
 Protein Accession #: NP_005039

1 11 21 31 41 51
 MAGLGASLHV WGNWMLGSL LARAQLSDG TITIEBQIV LKAKVQCEL NITAIQLQEG 60
 GNCFFPEWDG ICWPRGTVG ISAVPCPPYI YDFNHKGVA FRHCNPNGTW DFMHSLNKTW 120
 NYSDCLRFL QPDISIGKQE FFERLYVMYV VGYISIFGSL AVAILIIGYF RRLHCTRNHY 180
 RMHLFVSFML RATSIFVKDR VVHAHIGVKE LESLIMQDDP QNSIEATSV DSYQYICKIAV 240

WO 02/086443

PCT/US02/12476

5 VMFIYFLATN YYWILVEGLY LBNLIFVAFV SDTRYLWGFV LIGWGFPAAF VAAHVARAT 300
LADARCWELS AGDIKWIYQA PILAAIGLNF ILFLNTRVVL ATRKIWETNAV GHDTRKQYRK 360
LAKSTLVVL VFGVHYIVFV CLPHSPTGLG WEIRMHCELF FNSFQGFVVS IICYCNGEV 420
QAEVKMWSR WNLSDWKRT PPCGSRRCGS VLTVTVTHSTS SQSQVAASTR MVLISGKAAK 480
IASRQPDSDI TLPQYVWSNS EQDCLPHSFH EETKEDSGRQ GDDILMEKPS RPNESNPDTE 540
GCGGETEDVL

Seq ID NO: 664 DNA sequence
Nucleic Acid Accession #: NM_012152
Coding sequence: 43..1104

10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
1
11
21
31
41
51
CTTCTTAAAT TTTCTTCTCA GGATGTTTAC TTCTTCTCCA CAATGAATGA GTGTCACTAT 60
GACAAGCACA TGGACTTTTT TTATAATAGG AGCAACACTG ATACTGTGCA TGACTGGACA 120
GGAACAAGC TTGTGATGTG TTTGTGTGTT GGGACGTTTT TCTGCGTGT TATTTTTTTT 180
TCTAATCTCT TGGTCATCGC GGCAGTGATC AAAAACACAA AATTTCATTT CCCCCTCTAC 240
TACCTGTGGG CTAATTTAGC TGCTGCCGAT TTCTTCGCTG GAATTGCCTA TGTATTCTCG 300
ATGTTTAAAC CAGGCCCACT TTCAAAAAC TTGACTGTCA ACCGCTGGTT TCTCCGTCAG 360
GGGCTTCTCG ACAGTAGCTT GACTGCTTCC CTCACCAACT TGCTGGTTAT CGCCGTGGAG 420
AGGCACATGT CAATCATGAG GATGCGGGTC CATAGCAACC TGACCAAAA GAGGGTGACA 480
CTGCTCATTT TGCTGTCTG GGCATCGCCC ATTTTATAGG GGGCGTCCC CACACTGGGC 540
TGGAAATGCC TCTGCAACAT CTCTGCCTGC TCTTCCCTGG CCCCATTTA CAGCAGGAGT 600
TACCTGTGTT TCTGGACAGT GTCCAACTCT ATGGCCTTCC TCATCATGGT TGTGGTGTAC 660
CTGCGGATCT ACGTGTACGT CAAGAGGAAA ACCAACGTCT TGTCTCCGCA TACAAGTGGG 720
TCCATCAGCC CGCCGAGGAG ACCCATGAAG CTAATGAAGA CGGTGATGAC TGTCTTAGGG 780
GCGTTTGTGG TATGCTGGAC CCGGGGCTCG GTGGTCTGCG TCCTCGACGG CCTGAACTGC 840
AGGCAGTGTG GCGTGCAGCA TGTGAAAAGG TGGTTCCTGC TGCTGGCGCT GCTCAACTCC 900
GTGCTGAACC CACATCACTA CTCCTACAAG GACGAGGACA TGTATGGCAC CATGAAGAAG 960
ATGATCTGCT GCTTCTCTCA GGAGAACCCA GAGAGGCGTC CCTCTCGCAT CCCCCTCACA 1020
GTCCTCAGCA GAGTGACAC AGGCAGCCAG TACATAGAGG ATAGTATTAG CCAAGGTGCA 1080
GTCTGCAATA AAAGCACTTC CTAAACTCTG GATGCTCTCT GGGCCACCCA GGTGATGACT 1140
GTCTTAGG

Seq ID NO: 665 Protein sequence
Protein Accession #: NP_036284

1
11
21
31
41
51
MNECHYDKHM DFFYNRSNTD TVDDWTGTLK VIVLCVGTFF CLFIFFSNLS VIAAVIKNRK 60
FHPFFYYLLA NLAAADFFAG IAYVELMENT GPVSKLTIVN RWFLRQGLLD SSLTASLTNL 120
LVIAVERHMS IMRMVRHSNL TKKRVTLIL LWIAIAIFMG AVPTLGNCL CNISACSSLA 180
PIYSRSYLVE WTVSNLMAFL IMVVYLRIY VYVKRKTNLV SPHTSGSISR RRTPKLMKT 240
VMTVLGAFVV CWTPLVLL LDGLNCRQCG VQHVKRWFLL LALLNSVNP IISYKDEDM 300
YGTMKKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK STS

Seq ID NO: 666 DNA sequence
Nucleic Acid Accession #: NM_002821
Coding sequence: 150..3362

50
55
60
65
70
75
80
85
1
11
21
31
41
51
AACTCCCGCC TCGGAGCGCC TCGGGGTCGG GCTCCGCGTG CGGCTGCTGC TCGGGCGGCC 60
GGGCTCCGGT GCGTCCGCTT CCTGTGCCCG CGCGGAGACA GTCTGCGGCC CGCGGTGCGC 120
CCTCAGCTCC TTTTCTCTAG CCGCGCGCGA TGGGAGCTGC GCGGGGATCC CCGGCCAGAC 180
CCCGCCGGTT GCGCTCTGCT AGCGTCTCTG TGCTGCCGCT GCTGGGCGGT ACCCAGACAG 240
CCATTGTCTT CATCAAGCAG CGGTCTCTCC AGGATGCACT GCAGGGGCGC CGGGCGCTGC 300
TTGCTGTGTA GGTGTGAGGT CCGGGCCCGG TACATGTGTA CTGGCTGCTC GATGGGGCCC 360
CTGTCCAGAG CACGGAGCGG CGTTTCCGCC AGGGCAGCAG CCTGAGCTTT GCAGCTGTGG 420
ACCGGCTGCA GAGACTCTGG ACCTTCCAGT GTGTGGCTCG GGATGATGTC ACTGGAGAAG 480
AAGCCCGCAG TGCCCAAGCC TCCTTCAACA TCAAAATGGT TGAGGCAGGT CCGTGGTCTC 540
TGAAGCATCC AGCCTCGGAA GCTGAGATCC AGCCACAGAC CCAGGTACAC CTTCGTGCCC 600
ACATTGATGG GCACCTCGG CCCACCTACC AATGGTTCCG AGATGGGACC CCCCCTTCTG 660
ATGGTCAGAG CAACCAACAC GTCAGCAGCA AGGAGCGGAA CCTGACGCTC CGGCCAGCTG 720
GTCTGAGACA TAGTGGGCTG TATTCCTGCT GCGCCCAAG TGCTTTTGGC CAGGCTTGCA 780
GCAGCCAGAA CTTCACCTTG AGCATTGCTG ATGAAAGCTT TGCCAGGGTG GTGCTGGCAC 840
CCCAGGACGT GGTAGTAGCG AGGTATGAGG AGGCCATGTT CCATTGCCAG TTCTCAGCCC 900
AGCCACCCCG GAGCTCTGAG TGGCTCTTTG AGGATGAGAC TCCCATCACT AACCCAGTCT 960
GCCCCCACA CCTCCGAGA GCCACAGTGT TTGCCAACGG GTCTCTGCTG CTGACCCAGG 1020
TCCGCCACAG CAATGCAAGT ATCTACCGCT GCATTGGCCA GGGGCAGAGG GGGCCACCCA 1080
TCATCTCTGA AGCCACACTT CACCTAGCAG AGATTGAAGA CATGCCGCTA TTTGAGCCAC 1140
GGGTGTTTAC AGCTGGCAGC GAGGAGCGTG TGACCTGCCT TCCCCCAAG GGTCTGCCAG 1200
AGCCACAGCT GTGGTGGGAG CACGCGGGAG TCCGCTGCC CACCCATGCG AGGGTCTACC 1260
AGAAGGGCCA CGAGCTGGTG TTGGCCAATA TTGCTGAAAG TGATGCTGGT GTCTACACCT 1320
GCCAGCGGCG CAACCTGGCT GTGTCAGCGA GACAGGATGT CAACATCACT GTGGCCACTG 1380
TGCCCTCTCT GCTGAAGAAG CCCCAGACA GCCAGCTGGA GGAGGGCAAA CCGGCTACT 1440
TGGATTGCTT GACCCAGGCC ACACCAAAAC CTACAGTTGT CTGGTACAGA AACCATGTC 1500
TCATCTCAGA GGAATCAAGG TTGAGGTTCT TCAAGAAATG GACCTTGCGC ATCAACAGCG 1560
TGGAGGTGTA TGATGGGACA TGGTACCGTT GTATGAGCAG CACCCAGGCC GGCAGCATCG 1620
AGGCCCAAGC CCGTGTCCAA GTGCTGGAAA AGCTCAAGTT CACACCACA CCCCAGCCAC 1680
AGCAGTGCAT GGAGTTTTCAC AAGGAGGCA CGGTGCCCTG TTCAGCCACA GGGCCAGAGA 1740
AGCCCACTAT TAAGTGGGAA CCGGCAGATG GGAGCAGCCT CCAGAGTGG GTGACAGACA 1800
ACGCTGGGAC CCTGCATTTT GCCCGGGTGA CTCGAGATGA CGCTGGCAAC TACACTGCA 1860
TTGCTTCCAA CGGGCCGAGG GGCAGGATTC GTGCCCATGT CCAGCTCACT GTGGCAGTTT 1920
TTATCACTCT CAAAGTGAAA CCAGAGCGTA CGACTGTGTA CCAGGGCCAC ACAGCCCTAC 1980
TGCACTGCGA GGGCCAGGGG GACCCCAAGC CGCTGATTCA GTGAAAGGC AAGGACCGCA 2040
TCTTGAGGCC CACCAAGCTG GGACCCAGGA TGCACATCTT CCAGATGGC TCCTGGTGA 2100

WO 02/086443

PCT/US02/12476

5

10

15

20

25

30

35

TCCATGACGT GGCCTCTGAG GACTCAGGCC GCTACACCTG CATTGCAGGC AACAGCTGCA 2160
 ACATCAAGCA CACGGAGGCC CCCCTCTATG TCGTGGACAA GCCTGTGCCG GAGGAGTCGG 2220
 AGGGCCCTGG CAGCCCTGCC CCCTACAAGA TGATCCAGAC CATTGGGTGG TCGGTGGGTG 2280
 CCGCTGTGGC CTACATCATT GCGGTGCTGG GCCTCATGTT CTACTGCAAG AAGCGCTGCA 2340
 AAGCCAAGCG GCTCGAGAAG CAGCCCGAGG GCGAGGAGCC AGAGATGGAA TGCCTCAAGC 2400
 GAGGGCTCTT GCAGAACGGG CAGCCCTCAG CAGAGATCCA AGAAGAAAGT GCCTTGACCA 2460
 GCTTGGGCTC GGGCCCGCGG GCCACCAACA AACGCCACAG CACAAGTGAT AAGATGCACT 2520
 TCCCAAGGTC TAGCTGACAG CCCATCACC CGCTGGGGAA GAGTGAGTTT GGGGAGGTGT 2580
 TCCTGGCAAA GGCTCAGGGC TTGGAGGAGG GAGTGGCAGA GACCTCGTA CTGTGAAGA 2640
 GCCTGCAGAC GAAGGATGAG CAGCAGCAGC TGGACTTCGG GAGGGAGTTG GAGATGTTTG 2700
 GGAAGCTGAA CCACGCCAAC GTGGTGGCGG TCCTGGGGCT GTGCCGGGAG GCTGAGCCCC 2760
 ACTACATGGT GCTGGAATAT GTGGATCTGG GAGACCTCAA GCAGTTCCTG AGGATTTCCT 2820
 AGAGCAAGGA TGAAAAATTG AAGTCACAGC CCCTCAGCAC CAAGCAGAAG GTGGCCCTAT 2880
 GCACCCAGGT AGCCCTGGGC ATGGAGCACC TGTCCAACAA CCGCTTTGTG CATAAGGACT 2940
 TGGCTGCGCG TAACCTGCTG GTCACTGCCG AGAGACAAGT GAAGGTGTCT GCCCTGGGCC 3000
 TCAGCAAGGA TGTGTACAAC AGTGAGTACT ACCACTTCGG CCAGGCCTGG GTGCCGTGCG 3060
 GCTGGATGTC CCCGAGGAGG ATCCTGGAGG GTGACTTCTC TACCAAGTCT GATGCTGTGG 3120
 CCTTCGGTGT GCTGATGGG GAAAGTGTTA CACATGGAGA GATGCCCAT GGTGGGCAGG 3180
 CAGATGATGA AGTACTGGCA GATTGTCAGG CTGGGAAGGC TAGACTTCCT CAGCCGAGG 3240
 GCTGCCCTTC CAACACTATC CGGCTGATGC AGCGCTGCTG GGCCCTCAGC CCCAAGGACC 3300
 GGCCCTCTCT ATCAGGAGAT GCCAGCGCCC TGGGAGACAG CACCGTGAC AGCAAGCCGT 3360
 GAGGAGGGAG CCCGCTCAGG ATGGCTGGG CAGGGGAGGA CATCTTAGA GGGAGGCTCA 3420
 CAGCATGATG GGCAGATCC CTGTCTCTCT GGGCCCTGAG GTGCCCTAGT GCAACAGGCA 3480
 TTGCTGAGGT CTGAGCAGGG CCTGCTCTCT CCTCTCTCT CTACCCCTCA TCCTTTGGGA 3540
 GGCTGACTTG GACCCAACT GGGGACTAG GGCTTTGAGC TGGGCACTTT CCCTTGCCAC 3600
 CTCTCTCTCT ATCAGGAGCA GTGTGGGTGC CACAGGTAAC CCCAATTCTT GGCTTTCAAC 3660
 TTCTCCCTCT GACCGGGTCC AACTCTGCCA CTCTCTGCC AACTTTGCTT GGGGAGGGCT 3720
 AGGCTTGGGA TGAGCTGGGT TTGTGGGAG TTCTTAATA TTCTCAAGT CTGGGCACAC 3780
 AGGGTTAATG AGTCTCTTGC CCACTGGTCC ACTTGGGGGT CTAGACCAGG ATTATAGAGG 3840
 ACACAGCAAG TGAGTCTCTC CCACTCTGGG CTGTGTCACA CTGACCAGCA CCCACGCTCT 3900
 CCCACCTCTT CTCTCTCTT CTCTCTCTAA GTGCTTGGCA GATGAAGGAG TTTTCAGGAG 3960
 CTTTGTACAC TATATAAACG GCGCTTTTGT TATGCACCAC GGGCGGCTTT TATATGTAAT 4020
 TGCAGCGTGG GGTGGGTGGG CATGGGAGGT AGGGGTGGGC CCTGAGATG AGGAGGGTGG 4080
 GCCATCCTTA CCCACACTT TTTTGTGTGT CGTTTGTGT TTGTTTGTGT TTTTGTGTGT 4140
 TGTTTTGTGT TTTTCACTCG CTGCTCTCAA TAAATAAGCC TTTTTTA

Seq ID NO: 667 Protein sequence
Protein Accession #: NP_002812

40

45

50

55

60

1 11 21 31 41 51
 | | | | |
 MGAARGSPAR PRRLPLLSVL LPLLLGGTQT AIVFIKQPSS QDALQRRRL LRCEVEAPGP 60
 VHYVILLDGA PVQDTERREA QGSSLSFAAV DRLQDSGTFO CVARDDVTGE EARSANASFN 120
 IKWIEAGPVV LKHPASEAEI QPQTQVTLRC HIDGHPRTPT QWFRDGTPLS DGQSNHTVSS 180
 KERNLTLRPA GPBHSGLYSC CAHSAPGQAC SSQNFTLSIA DESFARVULA PQDVVVARYE 240
 EAMPHCQPSA QPPPSLQWLF EDETPIITNS RPHLRRAIV PANGSLLLTQ VRPRNAGIYR 300
 CIGQGQRGPP IILEATLHLA EIDEMPLFEP RVFTAGSEER VTCLPPKGLP EPSVNWHEAG 360
 VRLPTHGRVY QKGHELVLAN IAESDAGVYT CHAANLAGOR RQDVNITVAT VPSWLKPKPD 420
 SQLEEGKPGY LDCLTQATPK PTVVWYRNQM LISEDSRFEV FKNGLTRINS VEVYDGTWYR 480
 CMSSTPAGSI EAQARVQVLE KLKFTPPPPQ QCMFEDKEA TVPCSATGRE KPTIKWERAD 540
 GSSLPEWYTD NAGTCLHFAV TRDDAGNYTC IASNGPQGI RAHVQLTVAV FITPKVEPER 600
 TTVYQGHIAL LQCEAGGDPK PLIQWKGKDR ILDPKTLGPR MHIFONGSLV IHDVAPEDSG 660
 RYTCIAGNSC NIKHTEAPLY VVDKVPVEES EGPSPFPYK MIQTGLSVG AAVAYIIAIV 720
 GLMFYCKKRC KAKRLQKQPE GEEPEMECLN GGPLQNGQPS AEIQEVALT SLGSGPAATN 780
 KRHSTSDKMH FPRSLQPIIT TLGKSEFGEV FLAKAQGLEE GVAETLVLVK SLQTKDBQQQ 840
 LDFRLELMF KHLNHNANVVR LLGLCREAEP HYMVLEYVDL GDLKQFLRIS KSKDEKLKSO 900
 PLSTKQKVAL CTQVALGMEH LSNRFRVHKD LAARNCLVSA QRQVKVSALG LSKDVYNSEY 960
 YHFRQAWVPL RNMSPAILB GDFSTKSDVW AFGVLMWEVF THGEMPHGGQ ADEVDLADLQ 1020
 AGKARLPQPE GCPSKLYRLM QRCWALSPKD RPSFSEIASA LGDSTVDSPK

Seq ID NO: 668 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..1389

65

70

75

80

85

1 11 21 31 41 51
 | | | | |
 ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGA GAGATTAGA TGACAGAGAA 60
 ACCCTCTGTT CTGAACATGA GTATAAAGAG AAAACCTGTC AGTCTGCTGC TCTTTTAAAT 120
 GTTGTCAACT CGATTATAGG ATCTGGTATA ATAGGATTGC CTTATTCAAT GAAGCAAGCT 180
 GGGTTTCCTT TGGGAATATT GCTTTTATTC TGGGTTTCAT ATGTTACGGA CTTTCCCTT 240
 GTTTTATTGA TAAAAGGAGG GGCCCTCTCT GGAACAGATA CCTACCACTC TTTGGTCAAT 300
 AAAACTTTCC GCTTTCAGG GTATCTGCTC CTCTCTGTTT TCAGTTTCTT GTATCCTTTT 360
 ATAGCAATGA TAAGTTACAA TATAATAGCT GGAGATACTT TGAGCAAAGT TTTTCAAAGA 420
 ATCCCAAGAG TTGATCTGTA AAACGTGTTT ATTGGTGGCC ACTTCATTAT TGGACTTTCC 480
 ACAGTTACCT TTACTCTGCC TTTATCCTTG TACCGAAATA TAGCAAAGCT TGGAAAGGTC 540
 TCCTCATCT CTACAGGTTT AACAACTCTG ATTCTTGGAA TTGTAATGGC AAGGGCAATT 600
 TCACGTGGTG CACACATACC AAAAACAGAA GACGCTGGG TATTGCAAAA GCCCAATGCC 660
 ATTCAAGCGG TCGGGGTTAT GTCTTTTGCA TTTATTGCC ACCATACTC CTCTTAGTT 720
 TACAGTTCTC TAGAAGAACC CACAGTAGCT AAGTGGTCCC GCCTTATCCA TATGTCCATC 780
 GTGATTTCGT TATTATCTG TATATTCTTT GCTACATGTG GATACCTGAC ATTACTGGC 840
 TTCACCAAGG GGGACTTATT TGAAAATTAC TGCAGAAATG ATGACCTGGT AACATTGGGA 900
 AGATTGTTGT ATGGTGTGAC GTGCTTTTGT ACATACCTTA TGGAAATGCT TGTGACAAGA 960
 GAGGTAATTG CCAATGTGTT TTTTGGTGGG AATCTTTTCA CGGTTTCCA CATTGTTGTA 1020
 ACAGTGATGG TCATCACTGT AGCCACGCTT GTGTCAATGC TGATTGATTG CCTGGGGATA 1080
 GTTCTAGAAC TCAATGTGTT GCTCTGTGCA ACTCCCTCCA TTTTATCAT TCCATCAGCC 1140
 TGTATCTGTA AACTGTCTGA AGAACCAAGG ACACACTCCG ATAAGATTAT GTCTTGTGTC 1200
 ATGCTTCCCA TTGGTGTGTT GGTGATGTTT TTTGATTTCG TCATGGCTAT TACAAATACT 1260

WO 02/086443

CAAGACTGCA CCCATGGGCA GGAAATGTTT TACTGCTTC CTGACAATTT CTCTCTCACA 1320
 AATACCTCAG AGTCTCATGT TCAGCAGACA ACACAACCTT CTACTTTAAA TATTAGTATC 1380
 TTTCATGA

5 Seq ID NO: 669 Protein sequence
 Protein Accession #: Eos sequence

10 1 11 21 31 41 51
 MG YQRQEPVI PPQRDLDDRE TLVSEHEYKE KTCQSAALFN VVNSIIGSGI IGLPYSMKQA 60
 GFPLGILLF WVSVYVDFSL VLLIKGGALS GTDTYQSLVN KTFGFPGYLL LSVLQFLYPF 120
 IAMISYNIIA GDTLSKVFOR IPGVDPENVP IGRHFIIGLS TVTFTLPLSL YRNIKLGKV 180
 SLISTGLTTL ILGIVMARAI SLGPHIPKTE DAWVFAKPA IQAVGVMSFA FICHNSFLV 240
 YSSLEETVA KWSRLIHMSI VISVFICIFF ATCGYLTFTG FTQGDLPENY CRNDDLVTFG 300
 15 RFCYGVTVIL TYPMECFVTR EVIANVFFGG NLSSVPHIVV TVMVITVATL VSLIDCLGI 360
 VLELNGVLCA TPLIFIIPSA CYLKLSEEP THSDKIMSCV MLPIGAVVMV FGFVMAITNT 420
 QDCTHGQEMF YCFPDNFSLT NTSSEHVQQT TQLSTLNI SI FQ

20 Seq ID NO: 670 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..1284

25 1 11 21 31 41 51
 ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCGAGA GAGGATTGCC TTATTCAATG 60
 AAGCAAGCTG GGTTCCTCTT GGAATATTG CTTTATTCTT GGGTTTCATA TGTACAGAC 120
 TTTCCCTTG TTTTATGAT AAAAGGAGGG GCCCTCTCTG GAACAGATAC CTACCACTCT 180
 TTGGTCAATA AAATCTTCGG CTTTCCAGGG TATCTGCTCC TCTCTGTTCT TACAGTTTGT 240
 30 TATCCTTTTA TAGCAATGAT AAGTTACAAT ATAATAGCTG GAGATACCTT GAGCAAAAGT 300
 TTTCAAAGAA TCCAGGAGT TGATCCTGAA AACGTGTTTA TTGGTCGCCA CTTCATTATT 360
 GGACTTTCCA CAGTTACCTT TACTCTGCTT TATCTCTTGT ACCGAAATAT AGCAAAGCTT 420
 GGAAAGTCTT CCTCATCTC TACAGGTTTA ACAACTCTGA TTCTTGAAT TGTATGGCA 480
 AGGGCAATTT CACTGGGTCC ACACATACCA AAAACAGAAG ACGCTTGGGT ATTTGCAAG 540
 35 CCCAATGCCA TTCAGCGGT CGGGGTTATG TCTTTGTCAT TTATTGCCA CCATAACTCC 600
 TTCTTAGTTT ACAGTCTCT AGAAGAAGCC ACAGTAGCTA AGTGGTCCCG CCTTATCCAT 660
 ATGTCCATCG TGATTTCTGT ATTTATCTGT ATATTCTTTG CTACATGTGG ATACTTGACA 720
 TTTACTGGCT TCACCCAGG GGAATTTATG GAAATTAAT GCAGAAATGA TGACCTGGTA 780
 ACATTGGGAA GATTTGTGTA TGGTGTCACT GTCAATTTGA CATACCTTAT GGAATGCTTT 840
 40 GTGACAAGAG AGGTAATGTC CAATGTGTTT TTTGGTGGGA ATCTTTTCAT GGTTTTCCAC 900
 ATTGTTGTAA CAGTGATGGT CATCACTGTA GCCACGCTTG TGTCAATGCT GATTGATGTC 960
 CTCGGGATAG TTCTAGAACT CAATGGTGTG CTCTGTGCAA CTCCTTCAT TTTTATCAAT 1020
 CCATCAGCCT GTTATCTGAA ACTGTCTGAA GAACCAAGGA CACACTCCGA TAAGATTATG 1080
 TCTTGTGTC TGTCTCCCAT TGGTGTGTGT GTGATGGTTT TTGATTCGCT CATGGCTATT 1140
 45 ACAATACTC AAGACTGCAC CCATGGGCAG GAAATGTTCT ACTGCTTTCC TGACAAATTC 1200
 TCTCTCAAA ATACCTCAGA GTCTCATGTT CAGCAGACAA CACAACCTTC TACTTTAAAT 1260
 ATTAGTATCT TTCAACTCGA GTAA

50 Seq ID NO: 671 Protein sequence
 Protein Accession #: Eos sequence

55 1 11 21 31 41 51
 MG YQRQEPVI PPQRGLPYSM KQAGPPLGIL LFWVSVYVD PSLVLLIKGG ALSGTDYQS 60
 LVNKTGFPFG YLLSLVLQPL YPFIAMISYN IAGDTLSKV FORIPGVDP NVPFGRHFI 120
 GLSTVTPLP LSLYNIKAL GKVSLLTGL TLLILGIVMA RAISLGPHIP KTEDAWVFAK 180
 PNAIQAVGVV SFAFICHNS FLVYSSLEEP TVAKWSRLIH MSIVISVFIC IFATCGYLT 240
 FTGFTQGDLF ENYCRNDDL TFGRECYGVT VILTYMECF VTREIVANVF PGGNLSSVFH 300
 IVVTVMVITV ATLVSLLIDC LGIVLELNGV LCATPLIFII PSACYLKLSE EPRTHSDKIM 360
 60 SCVMLPIGAV VMVPGFVMAI TMTQDCTHGO EMFYCFPDNF SLTNTSESHV QOTTQLSTLN 420
 ISIFOLE

65 Seq ID NO: 672 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..1203

70 1 11 21 31 41 51
 ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCGAGT TTTCCCTTGT TTTATTGATA 60
 AAAGGAGGGG CCTCTCTGG AACAGATACC TACCAGTCTT TGGTCAATAA AACTTTCGGC 120
 TTTCCAGGGT ATCTGCTCCT CTCTGTTCTT CAGTTTTTGT ATCCTTTTAT AGCAATGATA 180
 AGTTACAATA TAATAGCTGG AGATACTTGG AGCAAGTTT: TTCAAGAAT CCCAGAGTT 240
 GATCTGAAA ACGTGTTTAT TGGTCCGCAC TTCAATTATG GACTTTCCAC AGTTACCTTT 300
 75 ACAGGTTTAA CAATCTGAT TCTTGAATT GTAATGGCAA GGGCAATTC ACTGGGTCCA 360
 CACATACCAA AAACAGAAGA OGCTGGGTA TTTGCAAGC CCATGCCAT TCAAGCGGTC 420
 GGGGTTATGT CTTTGGCATT TATTGGCCAC CATACTCCT TCTTAGTTTA CAGTTCTCTA 480
 GAAGAACCBA CAGTAGCTAA GTGGTCCCGC CTTATCCATA TGTCCATCGT GATTCTGTA 540
 TTTATCTGTA TATTCTTTGC TACATGTGGA TACTTGACAT TTACTGCTT CACCCAAGGG 600
 GACTTATTG AAAATCTGAT CAGAAATGAT GACCTGGTAA CATTTGGAAG ATTTGTTAT 660
 80 GGTGTCACTG TCAATTTGAC ATACCCTATG GAATGCTTTG TGACAAGAGA GGTAAATGCC 720
 AATGTGTTT: TTGGTGGGAA TCTTTCATCG GTTTCCACA TTGTGTAAAC AGTGATGGTC 780
 ATCACTGTAG CCACGCTGTG GTCAATGCTG ATTGATGCGC TCGGGATAGT TCTAGAATCT 840
 AATGGTGTGC TCTGTGCAAC TCCCTTCATT TTTATCATT CATCAGCCTG TTATCTGAAA 900
 CTGTCTGAAG AACCAAGGAC AACTCCGAT AAGATTATGT CTTGTGTGAT GCTTCCCAT 960
 85 GGTGCTGTGC TGTAGGTTT TGGATTGCTC ATGGCTATTA CAAATACTCA AGACTGCACC 1020
 CATGGGCAGG AAATGTTCTA CTGCTTTCCT GACAATTTCT CTCTCAGAAA TACCTCAGAG 1080
 TCTCATGTTT ACAGACACAC ACAACTTTCT ACTTTAAATA TTAGTATCTT TCAACTCGAG 1140
 1200

TAA

Seq ID NO: 673 Protein sequence
Protein Accession #: Eos sequence

5
10
15

1	11	21	31	41	51	
MGYQRQEPVI	PPQFSLVLLI	KGGALSGTDT	YQSLVNKTFG	FFGYLLLSVL	QPLYPFIAMI	60
SYNIIAGDTL	SKVFQRIQGV	DPENVFIGRH	FIIGLSTVTF	TLPLSLYRNI	AKLGKVSLSIS	120
TGLTTLILGI	VMARAIISLGP	HIPKTEDAWV	FAKPNAIQAV	GVMSFAPICH	HNSFLVYSSL	180
EEPTVAKWSR	LIHMSIVISV	PICIFFATCG	YLFTGTGTQ	DLFENYCRND	DLVTFGRFCY	240
GVTVILTYPM	ECFVTREVIA	NVFFGGNLSS	VFHIVVTVMV	ITVATLVSLI	IDCLGIVLEL	300
NGVLCATPLI	FIIPSACYLK	LSEEPRTSHD	KIMSCVMLPI	GAUVMVFGFV	MAITNTQDCT	360
HQQEMPYCFP	DNFSLINTSE	SHVQQTQLS	TLNISIFQLE			

Seq ID NO: 674 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..1140

20
25
30
35
40

1	11	21	31	41	51	
ATGGGCTACC	AGAGGCAGGA	GCCTGTCATC	CCGCCGCAGG	TCAATAAAC	TTTCGGCTTT	60
CCAGGGTATC	TGCTCCTCTC	TGTTCTTCAG	TTTTTGTATC	CTTTTATAGC	AATGATAAGT	120
TACAAATATA	TAGCTGGAGA	TACTTTGAGC	AAAGTTTTTC	AAAGAATCCC	AGGAGTTGAT	180
CCTGAAACAG	TGTTTATTGG	TCGCCACTTC	ATTATTGGAC	TTTCCACAGT	TACCTTTACT	240
CTGCCTTTAT	CCTTGTACCG	AAATATAGCA	AAGCTTGGAA	AGGTCTCCCT	CATCTCTACA	300
GGTTTAACAA	CTCTGATTCT	TGGAATTGTA	ATGGCAAGGG	CAATTTCAC	GGGTCCACAC	360
ATACCAAAAA	CAGAAGACGC	TGGGTATTTT	GCAAAGCCCA	ATGCCATTCA	AGCGGTCCGG	420
GTTATGTCCT	TTCATTATAT	TGCCCACCAT	AACCTCTTCT	TAGTTTACAG	TTCTCTAGAA	480
GAACCCACAG	TAGCTAAGTG	GTCCCGCCTT	ATCCATATGT	CCATCGTGAT	TTCTGTATTT	540
ATCTGTATAT	TCTTTGCTAC	ATGTGGATAC	TTGACATTTA	CTGGCTTCAC	CCAAGGGGAC	600
TTATTGAAA	ATTACTGCAG	AAATGATGAC	CTGGTAACAT	TTGGAAGATT	TTGTATGGT	660
GTCACGTGCA	TTTTCACATA	CCCTATGGAA	TGCTTTGTGA	CAAGAGAGGT	AATTGCCAAT	720
GTGTTTCTTG	GTGGGAATCT	TTCATCGGTT	TTCCACATTG	TTGTAACAGT	GATGGTCATC	780
ACTGTAGCCA	CGCTTGTGTC	ATTGCTGATT	GATTGCCTCG	GGATAGTTCT	AGAAGCTCAAT	840
GGTGTGCTCT	GTGCACTCC	CCTCATTTT	ATCATTCCAT	CAGCCTGTTA	TCTGAACTG	900
TCTGAAGAAC	CAAGGACACA	CTCCGATAAG	ATTATGTCTT	GTGTCATGCT	TCCCATGGT	960
GCTGTGTGTA	TGGTTTTGTG	ATTGCTCATG	GCTATTACAA	ATACTCAAGA	CTGCACCCAT	1020
GGGCAGGAAA	TGTTCTACTG	CTTCTCTGAC	AATTTCTCTC	TCACAAATAC	CTCAGAGTCT	1080
CATGTTGAGC	AGACAACACA	ACTTCTTACT	TAAATATTA	GTATCTTCA	ACTCGAGTAA	

Seq ID NO: 675 Protein sequence
Protein Accession #: Eos sequence

45
50
55

1	11	21	31	41	51	
MGYQRQEPVI	PPQVKNKTFG	PGYLLLSVLQ	FLYPFIAMIS	YNIAGDTLS	KVFQRIQGV	60
PENVFIGRHF	IIGLSTVFTF	LPLSLYRNIA	KLKGVSLIST	GLTTLILGIV	MARAIISLGP	120
IPKTEDAWV	AKPNAIQAVG	VMSFAPICHH	NSFLVYSSLE	EPTVAKWSRL	IHMSIVISVF	180
ICIFFATCGY	LTFGTGTQGD	LFENYCRNDD	LVTGFRFCYQ	VTIVILTYPME	CFVTVREVIAN	240
VFFGGNLSSV	FHIVVTVMVI	TVATLVSLLI	DCLGIVLELN	GVLCATPLIF	IIPSACYLKL	300
SEEPRTSHDK	IMSCVMLPIG	AVVMVFGFVM	AITNTQDCTH	GQEMPYCFPD	NFSLINTSES	360
HVQQTQLST	LNISIFQLE					

Seq ID NO: 676 DNA sequence
Nucleic Acid Accession #: NM_006853.1
Coding sequence: 26..874

60
65
70
75
80

1	11	21	31	41	51	
AGGAATCTGC	GCTCGGGTTC	CGCAGATGCA	GAGGTTGAGG	TGGCTGCGGG	ACTGGAAGTC	60
ATCGGGCAGA	GGTCTCACAG	CAGCCAAGGA	ACCTGGGGCC	CGCTCCTCCC	CCCTCCAGGC	120
CATGAGGATT	CTGCAGTTAA	TCTGCTTGC	TCTGCAACA	GGCCTGTAG	GGGAGAGAC	180
CAGGATCATC	AAGGGGTTCC	AGTGCAAGCC	TCACTCCAG	CCCTGGCAGG	CAGCCCTGTT	240
CGAGAAGACG	CGGTACTACT	GTGGGGCGAC	GCTCATCGCC	CCAGATGGC	TCTGACAGC	300
AGCCCACTGC	CTCAAGCCCC	GCTACATAGT	TCACCTGGGG	CAGCACAACC	TCCAGAAGGA	360
GGAGGGCTGT	GAGCAGACCC	GGACAGCCAC	TGAGTCCTTC	CCCCACCCCG	GCTTCAACAA	420
CAGCCTCCCC	AACAAGACC	ACCGCAATGA	CATCATGCTG	GTGAAGATGG	CATCGCCAGT	480
CTCCATCACC	TGGGCTGTGC	GACCCCTCAC	CCTCTCTCA	CGCTGTGTCA	CTGCTGCAC	540
CAGCTGCCTC	ATTTCGGGCT	GGGGCAGCAC	GTCCAGCCCC	CAGTTACGCC	TGCTTCACAC	600
CTTGCGATGC	GCCAACATCA	CCATCATTTA	GCACCAGAAG	TGTGAGAAGC	CCTACCCCGG	660
CAACATCACA	GACACCATGG	TGTGTGCCAG	CSTGCAGGAA	GGGGGCAAGG	ACTCCTGCCA	720
GGGTGACTCC	GGGGGCCCTC	TGCTCTGTAA	CCAGTCTCTT	CAAGGCATTA	TCTCTGGGG	780
CCAGGATCCG	TGTGCGATCA	CCGAAAGCC	TGCTGTCTAC	ACGAAAGTCT	GCAATATGTT	840
GGACTGGATC	CAGGAGACGA	TGAAGAACAA	TTAGACTGGA	CCCAACCCAC	ACAGCCCATC	900
ACCTCCATT	TCCACTTGGT	GTTTGGTTCC	TGTTCACTCT	GTTAATAAGA	AACCTTAAGC	960
CAAGACCTCT	TACGACACAT	CTTGGGCTCT	CCTGGACTAC	AGGAGATGCT	GTCATTAAT	1020
AATCAACCTG	GGGTTCGAAA	TCACTGAGAC	CTGATTCAAA	ATTCTGCCTT	GAAATATTGT	1080
GACTCTGGGA	ATGACAACAC	CTGGTTTGT	CTCTGTTGTA	TCCCCAGCCC	CAAAGACAGC	1140
TCCTGGCCAT	ATATCAAGST	TTCAATAAAT	ATTGTCTAAA	TGAGTG		

Seq ID NO: 677 Protein sequence
Protein Accession #: NP_006844.1

85

1	11	21	31	41	51	
MRILQLILLA	LATGLVGGET	RIIKGFCKP	HSQPWQAALF	EKTRLLCGAT	LIAPRWLLTA	60

AHCLKPRYIV HLGQHNQKE EGCEQTRTAT ESFPHPGFNN SLPNKDHRND IMLVKMASPV 120
SITWAVRPLT LSSRCVTAGT SCLISGWST SSPQLRLPHT LRCANITIE HOKCENAYPG 180
NITDTMVCAS VQEGGKDSQ GDSGGPLVCN QSLQGIISWG QDFCAITRKP GUYTKVKRYV 240
DHIQETMKNN

Seq ID NO: 678 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..933

10 1 11 21 31 41 51
ATGTGCAGCA ATGGACGGTG CATCCCGGGC GCCTGGCAGT GTGACGGGCT GCCTGACTGC 60
TTCGACAAGA GTGATGAGAA GGAGTGCCCC AAGGCTAAGT CGAAATGTGG CCGACCTTC 120
TTCCCTGTGT CCAGCGGCAT CCATTGCATC ATTGGTCGCT TCCGGTGCAA TGGGTTTGAG 180
15 GACTGTCCCG ACTGCAAGAA CGGCCTCTGT ATTGACAAGA GCTTCATCTG CGATGGACAG 240
GCCCGCTACC ACTGCAAGAA CGGCCTCTGT ATTGACAAGA GCTTCATCTG CGATGGACAG 300
AATAACTGTC AAGACAAACG TGATGAGGAA AGCTGTGAAA GTTCTCAAGA ACCCGCAGT 360
GGCAGGTGT TTGTGACTTC AGAGAACCAC CTGTGTGATT ACCCCAGCAT CACCTATGCC 420
ATCATCGCA GCTCGGTAT TTTGTGCTG GTGGTGGCCC TGCTGGCACT GGTCTTGAC 480
20 CACGAGCGA AGCGGAACAA CCTCATGACG CTGCCCGTGC ACCGGCTGCA GCACCCCTGTG 540
CTGCTGTCCC CTGCTGTGGT CCTGGACAC CCCCAACCAT GCAACGTGAC CTACAACTGC 600
AATAATGCA TCCAGTATGT GGCACGCGG GCGGAGCAGA ATGCGTCGGA AGTAGGCTCC 660
CACCCCTCTT ACTCCGAGGC CTGTGTGGAC CAGAGGCGCT CGTGTATGTA CCTTCCTCA 720
CCGCCCTACT CTCTGACAC GGAATCTCTG AACCAAGCGG ACCTGCCCCC CTACCCGCTCC 780
25 CGGTCCGGGA GTGCCAACAG TGCCAGCTCC CAGGAGCGCA GCAGCCTCTT GAGCGTGAA 840
GACACGAGC ACAGCCCGG GCAGCCTGCG CCCAGGAGG GCATCTGCTA GCCCAGGGAC 900
TCTGAGCCA GCCAGGGCAC TGAAGAAGTA TAA

Seq ID NO: 679 Protein sequence
Protein Accession #: Eos sequence

1 11 21 31 41 51
MCSNGRCIPG AWQCDGLPDC FDKSDEKECP KAKSKCGPTF FPCASGIHCI IGRFRNGFE 60
35 DCPDGSDEEN CTANPLLCST ARYHCKNGLC IDKSFICDQ MNOQDNDSEB SCESSQEPGS 120
QVFEVTSNQ LVVYPSITYA IIGSSVIFVL VVALLALVLH HQRKRNLMT LPVHRLQHPV 180
LLSRLVVLH PHHCNVYINV NNGIQYVASQ AEQNASEVGS PPSYSEALD QRPAYDLPP 240
40 PPSYSDTESL NQADLPFYRS RSGSANSASS QAASSLLSVE DTSHPQPG PQEGTAEPRD 300
SEPSQGTTEE

Seq ID NO: 680 DNA sequence
Nucleic Acid Accession #: S78203.1
Coding sequence: 1..2190

45 1 11 21 31 41 51
ATGAATCCTT TCCAGAAAAA TGAGTCCAAG GAAACTCTTT TTTCACCTGT CTCCATTGAA 60
GAGGTACCAC CTCGACCACC TAGCCCTCCA AAGAAGCCAT CTCGCAAAAT CTGTGGCTCC 120
50 AACTATCCAC TGAGCATTGC CTTCATTGTG GTGAATGAAT TCTGCGACG CTCTTCCTAT 180
TATGGAATGA AAGCTGTGCT GATCCTGTAT TTCTGTATT TCTGCACTG GAATGAAGAT 240
ACCTCCACAT CTATATACCA TGCCCTCAGC AGCCTCTGTT ATTCTACTCC CATCTCGGA 300
GCAGCCATTG CTGACTCGTG GTTGGGAAAA TTCAAGACAA TCATCTATCT TCCTTGGTG 360
TATGTGCTTG GCCATGTGAT CAAGTCCCTG GGTGCTTAC CAATACTGGG AGGACAAGTG 420
GTACACACAG TCCATCATT GATCGCGCTG AGTCTAATAG CTTTGGGGAC AGGAGGATC 480
55 AAACCCGTGT TGGCAGCTTT TGGTGGAGAC CAGTTTGAAG AAAACATGC AGAGGAACGG 540
ACTAGATACT TCTCAGTCTT CTACCTGTCC ATCAATGCAG GGAGCTGTAT TTCTACATT 600
ATCACACCCA TGCTGAGAGG AGATGTGCAA TGTTTTGAG AAGACTGCTA TGCATTGGCT 660
TTTGGAGTTC CAGGACTGCT CATGTGAATT GCACTTGTG TGTTTGCAAT GGGAGGCAAA 720
ATATACAATA AACCAACCCC TGAAGGAAAC ATAGTGGCTC AAGTTTCAAT ATGTATCTGG 780
60 TTTGTATTTT CCAATCGTTT CAAGAACCGT TCTGGAGACA TTCCAAAGCG ACAGCACTGG 840
CTAGACTGGG CAGCTGAGAA ATATCCAAAG CAGCTCATT TGGATGTAAG GGCATGACC 900
AGGGTACTAT TCCTTTATAT CCCATTGCCC ATGTCTGSG CTCTTTTGA TCAGCAGGGT 960
TCCAGATGGA CTTTGCAAGC CATCAGGATG AATAGGAATT TGGGGTTTT TGTGCTTCAG 1020
CGGACACAGA TGCAGTTCTT AAATCCCTTT CTGGTTCTTA TCTTCATCCC GTTGTGTGAC 1080
65 TTTGTCAATT ATGCTCTGTT CTCCAAGTGT GGAATTAAT TCTCATCAT TAGGAAAAATG 1140
GCTGTGTGTA TGATCTAGC GTGCTTGCA TTTGCACTG CGGAGCTGT AGAGATAAAA 1200
ATAAATGAAA TGGCCCCAGC CCAAGTCAGT GTGGGAAATG AAAACAATTC TCTGTGATA 1320
CTGGCAGATG ATGAGGTGAA GGTGACAGTG GTGGGAAATG AAAACAATTC TCTGTGATA 1380
GAGTCCATCA AATCCTTTCA GAAAAACCA CACTATTCCA AACTGCACCT TGAGCATTCT 1440
70 AGCCAGGATT TTCACTTCCA CCTGAAATAT CACAATTTGT CTCTCTACAC TGAGCATTCT 1500
GTGCAGGAGA AGAATCGGTA CAGTCTGTGC ATTCTGTAAG ATGGGAACAG TATCTCCAGC 1560
ATGATGTTAA AGGATACAGA AAGCAAAACA ACCAATGGGA TGACAACCGT GAGGTTTGT 1620
AACAATTTGC ATAAAGATGT CAACATCTCC CTGAGTACAG ATACCTCTCT CAATGTTGTT 1680
75 GAAGACTATG GTGTGTCTGC TTATAGAAT GTGCAAGAG GAGAATACCC TGCACTGAC 1740
TGTAAGACAG AAGATAAGAA CTTTCTCTG AATTGGGTC TTCTAGACTT TGGTGCAGCA 1800
TATCTGTTG TTATTACTAA TAACACCAAT CAGGGTCTTC AGGCCTGGAA GATTGAAGAC 1860
ATTCCAGCCA ACAAATGTG CATTGCGTGG CAGCTACCAC AATATGCCCT GGTACAGCT 1920
GGGAGGTCA GTTCTCTGT CACAGGTCTT GAGTTTTCTT ATTCTCAGG TCCCTCTAGC 1980
80 ATGAAATCTG TGCTCCAGC AGCTTGCTA TTGACAATT CAGTTGGGAA TATCATCTG 2040
CTGTGTGTG CACAGTTCAG TGGCCTGTA CAGTGGGCGG AATCATTTT GTTTCTCTG 2100
CTCTGCTGG TGATCTGCTT GATCTTCTC ATCATGGCT ACTACTATG TCCTGTAAG 2160
ACAGAGGATA TGGGGGTCC AGCAGATAAG CACATTCCTC ACATCCAGGG GAACATGATC
AAACTAGAGA CCAAGAAGAC AAAACTCTGA

85 Seq ID NO: 681 Protein sequence
Protein Accession #: AAB34388.1

1 11 21 31 41 51
| | | | |
MNFPQKNESE ETLFSPVSIE EVFPRPPSPF KKPSPTICGS NYPLSIAFIV VNEFCERFSY 60
YGMKAVLILY FLYFLHWNED TSTSIYHAFS SLCYFTPILG AALADSWLKG FKTIYILSLV 120
5 YVLGHEVIKSL GALPILGGQV VHTVLSLIGL SLIALGTGGI KPCVAAFQGD QPEEKHAER 180
TRYFSVYFSL INAGSLISTP ITPMLRGDVQ CFGEDCYALA FGVPGLLMVI ALVVFAMGSK 240
IYNKPPPEGN IVAQVFKCIW FAISNRFKNR SGDIKPKQHW LDWAAEKYKP QLIMDVKALT 300
RVLPYIPLP MFWALLDQGG SRWTLQAIRM NRNLGFFVLQ PDQMQLNPF LVLIFIPLFD 360
FVIYRLVSKC GINFSSLRKM AVGMILACLA FAVAAAVEIK INEMAPAQSG PQEVFLQVIN 420
10 LADDEVKQTV VGNENSLLI ESIKSFQKTP HYSKLHLKTK SQDFHFLKY HNLSLYTEHS 480
VQKNWYSLV IREDGNSISS MMVKDTESKT TNGMTTVRFV NTLHKDVNIS LSTDTSLNVG 540
EDYGVSAVRT VQGEYPVAVH CRTEDKNFSL NLGLLDFGAA YLFVITNNTN QGLQAWKIED 600
IPANKMSIAW QLPOVALVTA GEVMSFVTGL EFSYSQAPSS MKSVLQAAWL LTIAGVNIIV 660
15 LVVAQFSGLV QWAEFILFSC LLLVICLIFS IMGYIYVPVK TEDMRGPADK HTPHQGNMI 720
KLETKKTKL

Seq ID NO: 682 DNA sequence
Nucleic Acid Accession #: NM_016077.1
Coding sequence: 128..667

1 11 21 31 41 51
| | | | |
TCGCTTTGTG ATTCTTGATC CGGAACCTTG TCACCCAGGA ACCCGGGAAG AGGTAGCTCA 60
CGGATAGAAA ACGTGTTCGC TTGCCAGAA GAAGGGAAGG CGGAGTGAG GAAAGGAGGT 120
25 ACTGTAGATG CCTTCCAAAT CCTTGTATTG GAATATTTG GCTCATCCCA GTACACTCGG 180
CTTGCTCTGT GGAGTGTCTT GTGGCATGTG CCTGGGCTGG AGCCTTCGAG TATGCTTTGG 240
GATGCTCCTCC AAAAGCAAGA CGAGCAAGAC ACACACAGAT ACTGAAAGTG AAGCAAGCAT 300
CTTGGGAGAC AGCGGGGATG ACAAGATGAT TCTTGTGGTT CGAAATGACT TAAAGATGGG 360
AAAAGGAAA CTGGCTGCC AGTGCTCTCA TGCTGCTGTT TCAGCCTACA AGCAGATTCA 420
30 AAGAAGAAAT CCTGAAATGC TCAACAATG GGAATACTGT GGCAGCCCA AGTGTGTGTG 480
CAAAGCTCTT GATGAAGAAA CCTGTATTGC ATTATTGGCC CATGCAAAA TGCTGGGACT 540
GACTGTAAAT TTAATTCAG ATGCTGGACG TACTCAGATT GCACCAAGCT CTCAAACTGT 600
CCTAGGATG GGGCAGGAC CAGCAGACCT AATTGACAAA GTCACTGGTC ACCTAAACT 660
35 TTACTAGGTG GACTTTGATA TGACAACAAC CCTTCCATCA CAAGTGTGTT AAGCCTGTCA 720
GATTCTAACA ACAAAGCTG AATTTCTTCA CCCAATTAA ATGTTCTTGA GATGAAAATA 780
AAACCTATTC CCATGTTCTA AAAAAA

Seq ID NO: 683 Protein sequence.
Protein Accession #: NP_057161.1

1 11 21 31 41 51
| | | | |
MPSKSLVMEY LAHPSTLGLA VGVACGMCLG WSLRVCFGML PKSKTSKTHT DTESEASILG 60
DSGEYKMLLV VRNDLMKGK KVAACQSHAA VSAYKQIQRR NPMLKQWEY CGQPKVVVKA 120
45 PDEETLIALL AHAKMLGLTV SLIQDAGRTO IARGSQTVLG IGPFPADLID KVTGHLKLY

Seq ID NO: 684 DNA sequence
Nucleic Acid Accession #: NM_004864.1
Coding sequence: 26..952

1 11 21 31 41 51
| | | | |
CGGAACGAGG GCAACCTGCA CAGCCATGCC CGGCAAGAA CTCAGGACGG TGAATGGCTC 60
TCAGATGCTC CTGGTGTTC TGGTGTCTCT GTGGCTGCC CATGGGGGGC CCTGTCTCT 120
55 GGGCGAGGCG AGCGCGGCAA GTTCCCGGG ACCCTCAGAG TTGCACTCCG AAGACTCCAG 180
ATTCCGAGAG TTGCGGAAC GCTACGAGGA CCTGCTAACC AGGCTGCGGG CCAACCAGAG 240
CTGGGAAGAT TCGAACACCG ACCTGTGTCC GGGCCCTGCA GTCCGGATAC TCACGCCAGA 300
AGTGGGCTCG GAGTCCGGCG GCCACCTGCA CCTGGGTATC TCTGGGGCG CCTTCCCGA 360
60 GGGCTCTCCC GAGGCTCTCC GCTTCAACG GGCTCTGTTC CGGCTGTCCC CGAGGGGCTC 420
AAGGTCTGTG GAGTGACAC GACGCTGCG CGCTCAGCT AGCCTTGCAA GACCCCAAGC 480
GCCCGCGCTG CACTTCCGAC TGTGCGCGCC GCGGTGCGAG TCGGACCAAC TGCTGGCAGA 540
ATCTTCTGTC GCACGCGCCC AGCTGGAGTT GCATTTGCGG CCGCAAGCCG CCAGGGGGCG 600
CCGCAAGCG CGTCCGCGCA ACGGGGACGA CTGTCCGCTC GGGCCCGGGG GTTGTCTCCG 660
65 TCTGCACAG GTCCGCGCGT CGCTGGAAGA CCTGGGCTGG GCGGATGGG TGCTGTGCGC 720
ACGGGAGGTG CAAGTGACCA TGTGATCGG CGCGTGCCCG AGCCAGTTCC GGGCGGCAAA 780
CATGCACGCG CAGATCAAGA CGAGCCTGCA CCGCTGAAG CCGCACACGG AGCCAGCGCC 840
CTGCTCGCTG CCGCCAGCT ACAATCCCAT GGTGCTCAIT CAAAGACCG ACACGGGGGT 900
GTGCTCCAG ACCTATGATG ACTTGTTAGC CAAAGACTGC CACTGCATAT GAGCAGTCTC 960
70 GGTCTTCCA CTGTGACCT GCGGGGGGGA GCGGACCTCA GTTGTCTGCG CCTGTGGGAT 1020
GGGCTCAAGG TTCTGAGAC ACCGATTCC TGCCCAACA GCTGTATTTA TATAAGTCTG 1080
TTATTTATTA TTAATTTATT GGGGTGACCT TCTTGGGGAC TCGGGGGCTG GTCTGATGGA 1140
ACTGTGTATT TATTAAAC TCTGGTGATA AAAATAAAGC TGCTGAAC TTTAAAAA 1200
AAAA

Seq ID NO: 685 Protein sequence
Protein Accession #: NP_004855.1

1 11 21 31 41 51
| | | | |
MPGQELRTVN GSQMLLVLLV LSWLPHGGAL SLAEASRAST PGPSELHSED SRFRELKRY 60
EDLLRLRLAN QSWEDSNTDL VPAPAVRILT PEVRLSGGGH LHLRISRAAL PEGLEASRL 120
80 HRLRLRLSPT ASRSWDVTRP LRLQLSLARP QAPALHLRLS PPPSQSDQLL AESSSARPL 180
ELHLRPLQAR GRRRARARNG DDCPLGPRC RLHTRVRSAL EDLGNADWVL SPREVQVTMC 240
ICACPSQFRA ANMHAQIKTS LHLRLKPDTEP APCCVPASYN PMVLIQKTDI GVSILQTYDDL 300
LAKDCHCI

Seq ID NO: 686 DNA sequence

	1	11	21	31	41	51	
5	ACCAAATCAA	CCATAGGTCC	AAGAACAATT	GTCTCTGGAC	GGCAGCTATG	CGACTCACCG	60
	TGCTGTGTGC	TGTTGTGCTG	CTGCGCTGGCA	GCGTGGCCCT	CGCGCTGCCT	CAGGAGGCGG	120
	GGAGCAATGA	TGAGCTACAT	TGGGAACAGG	CTCAGGACTA	TCTCAAGAGA	TTTATCTCTT	180
10	ATGACTCAGA	AACAAAAAAT	GCCAACAGTT	TAGAAGCCAA	ACTCAAGGAG	ATGCACAAAAT	240
	TCTTTGGCCT	ACCTATAACT	GGAATGTTAA	ACTCCCGGCT	CATAGAAATA	ATGCAGAAGC	300
	CCAGATGTG	AGTTGCCAGT	GTTGCCAGAT	ACTCACTATT	TCCAAATAGC	CCAAAAATGGA	360
	CTTCCAAAGT	GGTCACCTAC	AGGATCGTAT	CATATATCTG	AGACTTACCG	CATATTATCG	420
	TGGATCGATT	AGTGTCAAAG	GCTTTAAACA	TGTGGGCGAA	AGAGATCCCC	CTGTACTTCA	480
	GGAAAGTTGT	ATGGGGAAAT	GCTGCACATC	TGATTGGCTT	TGCGCGAGGA	GCTCATGGGG	540
15	ACTCTACACC	ATTTGATGGG	CCAGGAAACA	GCGTGGCTGA	TGCTTTGGCG	CTCGGACAG	600
	GTCTCGGAGG	AGATGTCTAC	TTCGATGAGG	ATGAACCGTG	GACGGATGGT	AGCACTCTAG	660
	GGATTAACCT	CTCTGTATGT	GCAACTCATG	AACTTGGCCA	TTCTTTGGGT	ATGGGACAT	720
	CTCTGATCC	TATGTGAGCT	ATGATTCCAA	CCATATGGAA	TGGAGATCCC	CAAAATTTTA	780
20	AACTTTCCCA	GAGTGATATT	AAAGGCATCT	AGAAAACTTA	TGGAAGAGAGA	AGTAAATTCAA	840
	GAAGAAATA	GAACCTTCAG	GCAGAAATCC	CATTCTATCA	TTCTATGGAT	TTATATCAT	900
	TTTGGCACA	TCAGAAATGA	TAAGCACTGT	TCTTCACTC	CATTTAGCAA	TTATCTCACC	960
	CTTTTTTTAT	GCAGTGGT	TTTGAATGTC	TCTCACTCCT	TTTATGGT	AAATCTCTTT	1020
	ATGGTGTGAC	TGTGTTCTAT	TCCATCTATG	AGCTTTTGCA	GTGCGCGTAG	ATGTCATAAA	1080
	ATGTTACATA	CACAAATAAA	TAAATGTGTT	ATTCCATGGT	AAATTTA		
25							

	1	11	21	31	41	51	
30	MRLTVLCAVC	LLPGSLALPL	PQEAGGMSEL	QWEQAQDYLK	RPFLYDSETK	NANSLEAKLK	60
	EMQKFFGLPI	TGMLSNRVIE	IMQKPRCGPV	DAVIESLFPN	SPKMTSXCVT	YRIVSYTRLR	120
	PHITVDRVLS	KALANMWCKEI	PLPHKRCVVG	TDADIMIGPAR	GAHGDSYPPD	PGPDNLAHAF	180
	APGTLGGDA	HFGDEDERWT	KSSSLINFLY	AATHELGHSL	GMGHSSDPNA	VMYPTYNGGD	240
35	PONFKLSODD	IKEDIOKLYG	RNSNRKK				

40	1	11	21	31	41	51	
	ATGACAGGAG	TGTTTGACAG	AAGGGTCCCC	AGCATCCGAT	CGGGCGACTT	CCAAGCTCCG	60
	TTCCAGACGT	CGCAGCTAT	GACCATCATG	TCCTCAGAAAT	CGCCAACCTT	GCCCGAGCTCT	120
45	CTCAGCTACCG	ATTCTGACTA	CTACAGCCCT	ACGGGGGAGG	CCCGCCACGG	TACTGCTGCT	180
	CCTACCTCGG	TTCTCTATGG	CAAAGCTCTC	AACCCCTACC	AGTATCAGTA	CGCGGGGTG	240
	AACGCTCTCG	CGGGGAGGAT	CCCGACCAAA	GCTTATCGCC	ACTATAGCTA	CGCTAGCTCC	300
	TACCAACAGT	ACGGCGGCGC	CTACAAACCG	GTTCACAGAG	CACCAACAAC	GCCAGAGAAA	360
	GAAGTGACCG	ATCCGACAGT	GAGAATGGTG	AATGGCAAC	CAGAAGAAAT	TGGTAAACCC	420
50	AGGACTATTT	ATTCACGCTT	TCAAGCTGCC	GATTACAGCA	GAAAGTTTCA	GAGACTTCAG	480
	TACCTCGCTT	TGCGGGAAAG	CGCGCAGCTG	CGCCGCTCGC	TGGGATTGAC	CAAAACACAG	540
	GTGAAAATCT	GTTTTCAGAA	CAAAAGATCC	AAGATCAAGA	AGATCATGAA	AAACGGGGAG	600
	ATGCCCCCGG	AGCAGAGTCC	CAGCTCCAGC	GACCCAAATG	CGTGTAATCT	CGCGAGCTGT	660
	CTCCCGGTTG	GGGAGCCCCA	GGGCTCGTCC	CGCTCGCTCA	GCCACCAACC	TCATGCCCAT	720
55	CCTCGAGCTT	CAACACAGTC	CCACGAGTCC	AGGACTCTGG	AGGAACCTGC	ATCCTGGTAC	780
	ACAAGTGCAG	CCAGCTCAAT	CAATTCCCAC	CTGCGCCGCG	CGGGCTCCTT	ACAGCACCTG	840
	CTGGCGCTGG	CTCTCCGGAC	ACTCTATTAG				

	1	11	21	31	41	51	
65	MTGVFDRRP	SIRSGDFQAP	FQTSAA MHHP	SQESPTLPES	SATSDSYYP	TGGAPHGYCS	60
	PTSASYGKAL	NPYQYQHGV	NGGASGYPAK	AYADYSYASS	YHYQGYNQPK	VFSATNPQPK	120
	EVTSEPVRM	NGKPKVKVRK	RTIYSSPLQA	ALRRQFQTKQ	YLALPPEARAL	AASGLGTQTK	180
	VKIWFQNKRS	KIKKIMKNWE	MPPESSPSSS	DDMACNSPQS	PAVWPEQGSS	RLSLHHFPAH	240
	PPTSNSQSPAS	SYLSENSASPY	TYAASSINSH	LPPPGSLQHP	LALASGTLTY		

It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All publications, sequences of accession numbers, and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

- 1 1. A method of detecting a lung cancer-associated transcript in a cell
2 from a patient, the method comprising contacting a biological sample from the patient with a
3 polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence
4 as shown in Tables 1A-16.
- 1 2. The method of claim 1, wherein the polynucleotide selectively
2 hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-16.
- 1 3. The method of claim 1, wherein the biological sample is a tissue
2 sample.
- 1 4. The method of claim 1, wherein the biological sample comprises
2 isolated nucleic acids.
- 1 5. The method of claim 4, wherein the nucleic acids are mRNA.
- 1 6. The method of claim 4, further comprising the step of amplifying
2 nucleic acids before the step of contacting the biological sample with the polynucleotide.
- 1 7. The method of claim 1, wherein the polynucleotide comprises a
2 sequence as shown in Tables 1A-16.
- 1 8. The method of claim 1, wherein the polynucleotide is labeled.
- 1 9. The method of claim 8, wherein the label is a fluorescent label.
- 1 10. The method of claim 1, wherein the polynucleotide is immobilized on
2 a solid surface.
- 1 11. The method of claim 1, wherein the patient is undergoing a therapeutic
2 regimen to treat lung cancer.
- 1 12. The method of claim 1, wherein the patient is suspected of having lung
2 cancer.
- 1 13. A method of monitoring the efficacy of a therapeutic treatment of lung
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic
4 treatment; and

5 (ii) determining the level of a lung cancer-associated transcript in the
6 biological sample by contacting the biological sample with a polynucleotide that selectively
7 hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16,
8 thereby monitoring the efficacy of the therapy.

1 14. The method of claim 13, further comprising the step of: (iii) comparing
2 the level of the lung cancer-associated transcript to a level of the lung cancer-associated
3 transcript in a biological sample from the patient prior to, or earlier in, the therapeutic
4 treatment.

1 15. The method of claim 13, wherein the patient is a human.

1 16. A method of monitoring the efficacy of a therapeutic treatment of lung
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic
4 treatment; and

5 (ii) determining the level of a lung cancer-associated antibody in the biological
6 sample by contacting the biological sample with a polypeptide encoded by a polynucleotide
7 that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in
8 Tables 1A-16, wherein the polypeptide specifically binds to the lung cancer-associated
9 antibody, thereby monitoring the efficacy of the therapy.

1 17. The method of claim 16, further comprising the step of: (iii) comparing
2 the level of the lung cancer-associated antibody to a level of the lung cancer-associated
3 antibody in a biological sample from the patient prior to, or earlier in, the therapeutic
4 treatment.

1 18. The method of claim 16, wherein the patient is a human.

1 19. A method of monitoring the efficacy of a therapeutic treatment of lung
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic
4 treatment; and

5 (ii) determining the level of a lung cancer-associated polypeptide in the
6 biological sample by contacting the biological sample with an antibody, wherein the antibody
7 specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to
8 a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby
9 monitoring the efficacy of the therapy.

1 20. The method of claim 19, further comprising the step of: (iii) comparing
2 the level of the lung cancer-associated polypeptide to a level of the lung cancer-associated
3 polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic
4 treatment.

1 21. The method of claim 19, wherein the patient is a human.

1 22. An isolated nucleic acid molecule consisting of a polynucleotide
2 sequence as shown in Tables 1A-16.

1 23. The nucleic acid molecule of claim 22, which is labeled.

1 24. The nucleic acid of claim 23, wherein the label is a fluorescent label

1 25. An expression vector comprising the nucleic acid of claim 22.

1 26. A host cell comprising the expression vector of claim 25.

1 27. An isolated polypeptide which is encoded by a nucleic acid molecule
2 having polynucleotide sequence as shown in Tables 1A-16.

1 28. An antibody that specifically binds a polypeptide of claim 27.

1 29. The antibody of claim 28, further conjugated to an effector component.

1 30. The antibody of claim 29, wherein the effector component is a
2 fluorescent label.

1 31. The antibody of claim 29, wherein the effector component is a
2 radioisotope or a cytotoxic chemical.

1 32. The antibody of claim 29, which is an antibody fragment.

- 1 33. The antibody of claim 29, which is a humanized antibody
- 1 34. A method of detecting a lung cancer cell in a biological sample from a
2 patient, the method comprising contacting the biological sample with an antibody of claim
3 28.
- 1 35. The method of claim 34, wherein the antibody is further conjugated to
2 an effector component.
- 1 36. The method of claim 35, wherein the effector component is a
2 fluorescent label.
- 1 37. A method of detecting antibodies specific to lung cancer in a patient,
2 the method comprising contacting a biological sample from the patient with a polypeptide
3 encoded by a nucleic acid comprises a sequence from Tables 1A-16.
- 1 38. A method for identifying a compound that modulates a lung cancer-
2 associated polypeptide, the method comprising the steps of:
3 (i) contacting the compound with a lung cancer-associated polypeptide, the
4 polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least
5 80% identical to a sequence as shown in Tables 1A-16; and
6 (ii) determining the functional effect of the compound upon the polypeptide.
- 1 39. The method of claim 38, wherein the functional effect is a physical
2 effect.
- 1 40. The method of claim 38, wherein the functional effect is a chemical
2 effect.
- 1 41. The method of claim 38, wherein the polypeptide is expressed in a
2 eukaryotic host cell or cell membrane.
- 1 42. The method of claim 38, wherein the functional effect is determined by
2 measuring ligand binding to the polypeptide.
- 1 43. The method of claim 38, wherein the polypeptide is recombinant.

1 44. A method of inhibiting proliferation of a lung cancer-associated cell to
2 treat lung cancer in a patient, the method comprising the step of administering to the subject a
3 therapeutically effective amount of a compound identified using the method of claim 38.

1 45. The method of claim 44, wherein the compound is an antibody.

1 46. The method of claim 45, wherein the patient is a human.

1 47. A drug screening assay comprising the steps of
2 (i) administering a test compound to a mammal having lung cancer or a cell
3 isolated therefrom;
4 (ii) comparing the level of gene expression of a polynucleotide that selectively
5 hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16 in a
6 treated cell or mammal with the level of gene expression of the polynucleotide in a control
7 cell or mammal, wherein a test compound that modulates the level of expression of the
8 polynucleotide is a candidate for the treatment of lung cancer.

1 48. The assay of claim 47, wherein the control is a mammal with lung
2 cancer or a cell therefrom that has not been treated with the test compound.

1 49. The assay of claim 47, wherein the control is a normal cell or mammal.

1 50. A method for treating a mammal having lung cancer comprising
2 administering a compound identified by the assay of claim 47.

1 51. A pharmaceutical composition for treating a mammal having lung
2 cancer, the composition comprising a compound identified by the assay of claim 47 and a
3 physiologically acceptable excipient.

REVISED VERSION

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number
WO 2002/086443 A2

(51) International Patent Classification⁷: **C07H 21/02**,
21/04, C12Q 1/68

(21) International Application Number:
PCT/US2002/012476

(22) International Filing Date: 18 April 2002 (18.04.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/284,770 18 April 2001 (18.04.2001) US
60/290,492 10 May 2001 (10.05.2001) US
60/339,245 9 November 2001 (09.11.2001) US
60/350,666 13 November 2001 (13.11.2001) US
60/334,370 29 November 2001 (29.11.2001) US
60/372,246 12 April 2002 (12.04.2002) US

(71) Applicant (for all designated States except US): **PRO-
TEIN DESIGN LABS, INC** [US/US]; 225A Gateway
Boulevard, South San Francisco, CA 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **AZIZ, Natasha**
[US/US]; 411 California Avenue, Palo Alto, CA 94306
(US). **MURRAY, Richard** [US/US]; 22643 Woodbridge
Court, Cupertino, CA 95014 (US).

(74) Agent: **HALLUIN, Albert, P.; HOWREY SIMON
ARNOLD & WHITE LLP**, 301 Ravenswood Avenue, Box
34, Menlo Park, CA 94025 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Declaration under Rule 4.17:
— of inventorship (Rule 4.17(iv)) for US only

Published:
— with declaration under Article 17(2)(a); without abstract;
title not checked by the International Searching Authority

(48) Date of publication of this revised version: 17 June 2004

(15) Information about Correction:
see PCT Gazette No. 25/2004 of 17 June 2004, Section II.

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MOD-
ULATORS OF LUNG CANCER

(57) Abstract:

WO 2002/086443 A2

PATENT COOPERATION TREATY

PCT

DECLARATION OF NON-ESTABLISHMENT OF INTERNATIONAL SEARCH REPORT

(PCT Article 17(2)(a), Rule 13ter.1(c) and 39)

Applicant's or agent's file reference 18501-15-3PC	IMPORTANT DECLARATION	Date of mailing (day/month/year) 15 AUG 2003
International application No. PCT/US02/12476	International filing date (day/month/year) 18 April 2002 (18.04.2002)	(Earliest) Priority date (day/month/year) 10 May 2001 (10.05.2001)
International Patent Classification (IPC) or both national classification and IPC IPC(7): C07H 21/02, 21/04; C12Q 1/68 and US Cl.: 435/6, 536/23.1, 23.5		
Applicant EOS BIOTECHNOLOGY, INC		

This International Searching Authority hereby declares, according to Article 17(2)(a), that no international search report will be established on the international application for the reasons indicated below.

1. ☐ The subject matter of the international application relates to:
 - a. ☐ scientific theories.
 - b. ☐ mathematical theories
 - c. ☐ plant varieties.
 - d. ☐ animal varieties.
 - e. ☐ essential biological processes for the production of plants and animals, other than microbiological processes and the products of such processes.
 - f. ☐ schemes, rules or methods of doing business.
 - g. ☐ schemes, rules or methods of performing purely mental acts.
 - h. ☐ schemes, rules or methods of playing games.
 - i. ☐ methods for treatment of the human body by surgery or therapy.
 - j. ☐ methods for treatment of the animal body by surgery or therapy.
 - k. ☐ diagnostic methods practised on the human or animal body.
 - l. ☐ mere presentations of information.
 - m. ☐ computer programs for which this International Searching Authority is not equipped to search prior art.
2. ☐ The failure of the following parts of the international application to comply with prescribed requirements prevents a meaningful search from being carried out:

☐ the description
☐ the claims
☐ the drawings
3. ☒ The failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions prevents a meaningful search from being carried out:

☐ the written form has not been furnished or does not comply with the standard.
☒ the computer readable form has not been furnished or does not comply with the standard.
4. Further comments:

Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703)305-3230	Authorized officer <div style="text-align: center;"> CARLA J. MYERS PRIMARY EXAMINER </div> Telephone No. 703-308-0196
--	---

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To:
TOWNSEND AND TOWNSEND AND CREW LLP
TWO EMBARCADERO CENTER
EIGHTH FLOOR
SAN FRANCISCO, CA 94111-3834

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT OR THE DECLARATION

(PCT Rule 44.1)

Applicant's or agent's file reference 18501-15-3PC	Date of Mailing (day/month/year) 15 AUG 2003
International application No. PCT/US02/12476	International filing date (day/month/year) 18 April 2002 (18.04.2002)
Applicant EOS BIOTECHNOLOGY, INC	

1. ☐ The applicant is hereby notified that the international search report has been established and is transmitted herewith.

Filing of amendments and statement under Article 19:

The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

When? The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.

Where? Directly to the International Bureau of WIPO, 34, chemin des Colombettes
1211 Geneva 20, Switzerland, Facsimile No.: (41-22) 740.14.35

For more detailed instructions, see the notes on the accompanying sheet.

2. ☒ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect is transmitted herewith.

3. ☐ With regard to the protest against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

- ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.
- ☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. Reminders

Shortly after 18 months from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90 bis.1 and 90 bis.3, respectively, before the completion of the technical preparations for international publication.

Within 19 months from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later); otherwise the applicant must, within 20 months from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

In respect of other designated Offices, the time limit of 30 months (or later) will apply even if no demand is filed within 19 months.

See the Annex to Form PCT/IB/301 and, for details about the applicable time limits, Office by Office, see the *PCT Applicant's Guide*, Volume II, National Chapters and the WIPO Internet site.

Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703)305-3230	Authorized officer <i>Valerie Bell-Harris</i> Carla Myers Telephone No. 703-308-0196
--	---

Form PCT/ISA/220 (April 2002)

(See notes on accompanying sheet)